THE UNIVERSITY *of* EDINBURGH

Maintenance of Genomic Imprinting by G9a/GLP Complex of Histone Methyltransferases in Embryonic Stem (ES) Cells

Tuo Zhang

Thesis presented for the degree of Doctor of Philosophy

The University of Edinburgh

2014

Declaration

I declare that this thesis was composed by myself. The research in this thesis is my own, unless otherwise stated and has not been submitted for any other degree or professional qualification.

Tuo Zhang (August 2014)

Table of Contents

Index of Figures

Acknowledgements

To begin with, I would like to thank my supervisor, Dr. Irina Stancheva, who gave me the opportunity to join her lab and supervised me during the past four years in the University of Edinburgh. During my PhD, she not only taught and improved my experimental skills, more importantly she also encouraged and inspired me to improve my abilities of critical thinking and problem solving and also my English. I really appreciate for her precious suggestions throughout my PhD project, for the numerous times of the interesting discussion and for her critical reading and correcting of the draft of my thesis.

In addition, I also would like to thank my second supervisor Prof. Sir. Adrian Bird, my senior academics Prof. David Finnegan and Dr. Donald Macleod for their priceless advice and encouragement. Moreover, I really enjoyed the time that I worked with the past and current members in the Stancheva's lab, including Chao Li, Ausma Termanis, Joe Burrage, Katrina Gordon, Sadie Kemp, Simon Varzandeh and Burak Ozkan and thank for their help as well. Last, I would like to thank all the members in the Bird's Lab, especially Sabine Lagger, Martha Koerner and Elizabeth Wachter.

Furthermore, I would like to thank my farther, Prof. Zhiquan Zhang, for his patient education and enlightenments since I was a child. I also would like to thank my grandma, Ms. Jianyi Yang, my grandpa, Mr. Lianhai Zhang, and my mother Ms. Xiuhua Ren for their endless love and attention. Additionally, I would like to thank Miss. Xun Bao for her valuable encouragements in work and meticulous care in life.

Finally, I would like to thank the Darwin Trust of Edinburgh, which funded

my work and supported my living in the UK. Prof. Sir. Kenneth Murray and Prof. Noreen Murray, who created and founded the Darwin Trust, will forever in my heart and guide me to dedicate my entire life in biology!

Abstract

DNA methylation refers to an addition of a methyl group to the 5 position of the cytosine pyrimidine ring. As the best characterized epigenetic mark, DNA methylation plays an important role in a plethora of biological functions, including gene repression, genomic imprinting, silencing of retro-transposons and X chromosome inactivation. Genomic imprinting refers to the mono-allelic expression of certain genes according to their parent-of-origin. In mammals, the expression of imprinted genes is controlled by the *cis*-acting regulatory elements, termed imprinted control regions (ICRs). ICRs are marked by parent-of-origin-specific DNA methylation and loss of DNA methylation at ICRs also causes aberrant expression of imprinted genes. Therefore it is believed that the genomic imprinting is a DNA methylation-associated epigenetic phenomenon. As accurate expression of imprinted genes is essential for normal embryonic growth, energy homeostasis, development of the brain and behaviour and abnormal expression of imprinted genes leads to numerous clinical phenotype and human disorders, it is important to investigate how the imprinted DNA methylation is stably maintained in mammals.

DNA methyltransferases (DNMTs) are the main enzymes that play a in the establishment and maintenance of imprinted DNA methylation. In primordial germ cells (PGCs), DNMT3A and DNMT3L are involved in the establishment of imprinted DNA methylation. Whereas once established, the imprinted DNA methylation is maintained by DNMT1, DNMT3A and DNMT3B, but mainly by DNMT1. In addition, some other enzymes and DNA binding proteins also play a role in this process. One of the best examples is ZFP57, which forms a complex with

KAP1 and SETDB1. ZFP57 maintains imprinted DNA methylation by recognizing a methylated hexa-nucleotide and recruits DNMTs to the ICRs in mammalian embryonic stem (ES) cells. Interestingly, DNA methylation analysis combined with promoter microarrays carried out in our lab suggested that imprinted DNA methylation is absent from some of the maternal ICRs in ES cells genetically null for G9a, a histone H3 lysine 9 methylase. This indicates that G9a might also play a role in the maintenance of imprinted DNA methylation.

In my work, I found that the repressive H3K9me2 and imprinted DNA methylation are absent from several analysed ICRs in embryonic stem (ES) cells genetically null for either G9a or its partner histone methyltransferase GLP. A knockdown of G9a in ES cells reproduced these observations suggesting that G9a/GLP complex is required for the maintenance of imprinted DNA methylation. I also found that neither wild type nor catalytically inactive G9a can restore the loss of imprinted DNA methylation in *G9a-/-* ES cells. Chromatin immunoprecipitation (ChIP) combined with bisulfite DNA sequencing showed that imprinted DNA methylation was present on the H3K9me2-marked allele indicating a direct role for G9a in maintenance of genomic imprinting. Using a pharmacological inhibitor of G9a and mutagenesis analyses, I found that G9a maintains the imprinted DNA methylation independently of its catalytic activity and recruits DNMTs to the ICRs via its ankyrin repeat domain. Dimerization of G9a with GLP is also essential for the maintenance of genomic imprinting in ES cells. In summary, in addition to establish H3K9me2, histone methyltransferases G9a and GLP also play an essential role in the maintenance of genomic methylation imprints in ES cells.

Chapter 1 Introduction

1.1 Epigenetics

1.1.1 Genetics and epigenetics

In the mid-19th century, by tracing the phenotypic patterns in pea plants statistically, Gregor Mendel observed that organisms could inherit their traits in a discrete way. It is this cornerstone observation initiated the era of Mendelian genetics. Following Mendel's discoveries, scientists tried to determine which molecule is essential for the genetic inheritance. In 1911, due to the observation of a sex-linked white-eye mutation in *Drosophila*, Thomas Hunt Morgan revealed that genes are located on chromosomes. However, which macromolecule severs as the genetic material and is essential for the inheritance was still unclear.

Inspired by the discoveries of transformation phenomenon, Alfred Hershey and Martha Chase labeled the protein capsule and DNA of bacteriophages separately. They then performed an infectious assay and concluded that in eukaryotes, DNA is the essential genetic material, which is responsible for the inheritance (Hershey and Chase, 1952). However, it was still unclear why the differentiated tissue-specific cells, which contain identical genetic information, appear to have distinct morphologies and exert different physiological functions. Actually, it is this basic question that has advanced the understanding of epigenetics, a science about the genetic mechanisms beyond or upon the alteration of DNA sequences.

The history of epigenetics is associated with the study of evolution and developmental biology. Since the late 19th century, the functions of nature and nurture in the determination of organism development became a topic of hot debates. Although

DNA was confirmed to play a crucial role in genetics (Avery et al., 1944; Hershey and Chase, 1952), emerging evidence could not be explained by the classic genetic theory. For instance, in *Drosophila*, it was found that the eye colour variegates when rearrangements juxtapose the *white* gene with heterochromatin regions. In addition, the patterns of cilia in ciliates have also been found vary among individuals and are inherited clonally (Beisson and Sonneborn, 1965). These phenomena indicated that apart from genetics, some other mechanisms might also exist and together with the classic genetic mechanisms control the phenotype of the organism.

The original definition of epigenetics is "the study of heritable changes in the gene function, which cannot be explained by the changes of the genetic sequences" (Bird, 2002). Recently, Adrian Bird refined the definition of epigenetics as "the structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states" (Bird, 2007). In general, the mechanisms of epigenetics operate via two types of modifications: DNA methylation and post-translational histone modifications (PTHMs). In addition, since the structural remodelling of chromatin always coincides with the changes of histone modifications, chromatin remodelling is also considered as another aspect of epigenetic regulation. Today, emerging evidence has shown that epigenetic mechanisms play a role in a great number of biological processes, including gene and retrotransposon regulation, genomic imprinting, X chromosome inactivation and maternal effects (Bird, 2002; Jones, 2012).

1.1.2 Molecular mechanisms of epigenetics

Epigenetic marks refer to post-replication nucleic acid or post-translational protein modifications, which could influence the status of certain genes without altering the DNA sequences. The main epigenetic marks in mammals include DNA methylation

and post-translational histone modifications. Traditionally, DNA methylation is suggested to be associated with gene silencing. However, recent observations have shown that DNA methylation present at the gene body regions might facilitate transcription (Jones, 2012). Compared with DNA methylation, histone modifications could play either the positive or negative role in the regulation of gene expression. For example, H3K4me3 and H3K36me3, which present at promoters and the gene body regions respectively, play a role in genes activation, whereas some other histone modifications are associated with gene repression, including H3K9me3, H3K27me3 and H4K20me1.

Since to adapt to the different biological conditions, the expression of particular genes is dynamically regulated, it is undoubted that the deposition and removal of epigenetic marks are also dynamic, especially in the early developmental stage in mammals. However, most of the changes of epigenetic marks occurred within the course of an individual organism lifespan and only those occurred in the gametes might be transmitted to the next generation.

1.2 DNA methylation

In animals, DNA methylation refers to an addition of a methyl group to the 5-position carbon within the cytosine pyrimidine ring (Figure 1.1A). Although DNA methylation was originally uncovered approximately five decades ago (Bird, 2002), the underlying functions behind this mark were still obscure until the late of 1970s. Nowadays, DNA methylation has been found to play a pivotal role in many biological processes, including controlling of gene expression, X chromosome inactivation, maintenance of genome stability, regulation of genomic imprinting and even in the normal development of embryos and cell differentiation (Bird, 2002; Jones, 2012).

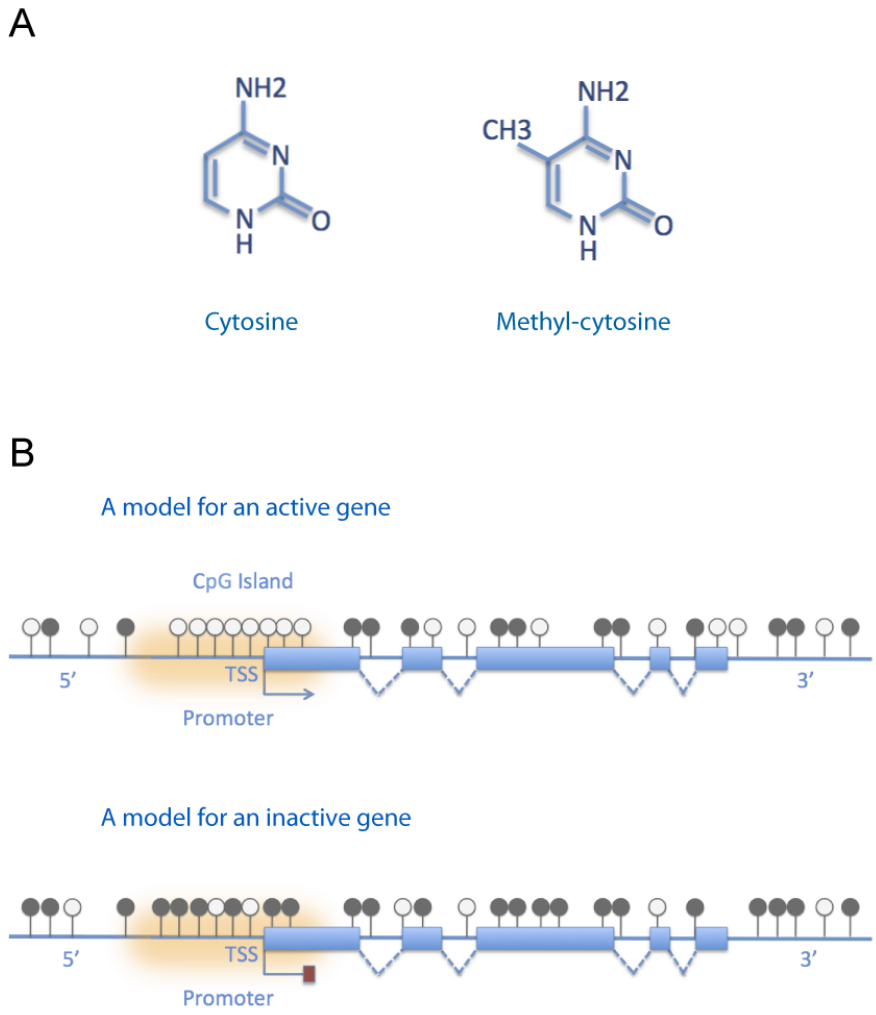


Figure 1.1 Structure and functions of DNA methylation in gene regulation.

(A) The structures of cytosine and 5'-methyl-cytosine. (B) The schematic models of the active and inactive genes. In general, the gene that contains a CpG island at its promoter is transcribed (top panel), whereas a gene containing DNA methylation at its promoter is nornally silenced. The white and black lollypops represent for the unmethylated and the methylated cytosine, respectively. The yellow shadow represents for the promoter, whereas TSS stands for the transcription start site.

1.2.1 Distribution of DNA methylation

Knowing how the DNA methylation is distributed across the genome is essential for understanding its function(s). The landscapes of DNA methylation in flowering plants and mammals are similar but nuanced. Actually among all the eukaryotes, plants contain the highest level of DNA methylation, with up to 50% of all the cytosine methylated in some species (Montero et al., 1992). In maize, the high frequency of methylated DNA in the genome is thought to be associated with extremely large numbers of transposons (Palmer et al., 2003; SanMiguel et al., 1996). In mammals, the proportion of the methylated cytosine within the genome is less than the proportion in plants. Take mouse for example, there is only 8% of the total cytosines are methylated. The majority of the methylated cytosines in mammals occur in the context of CpG dinucleotides, whereas in the context of CpHpG or CpHpH, where H represents a non-guanine nucleotide, the cytosines can also be methylated in plants (Bird, 2002; Bird, 1986). However, in eukaryotes, there is no DNA methylation in some species, such as *Saccharomyces cerevisiae* and *Drosophila*.

In plants, the vast majority of studies on DNA methylation were carried out in *Arabidopsis thaliana*, in which the distribution of DNA methylation displays as discrete islands, with all transposons and at least 20% of protein coding genes being methylated (Tran et al., 2005). Additionally, in *Arabidopsis thaliana*, the DNA methylation is preferentially biased away from promoters and the 3' ends of the protein coding genes. Since distribution of DNA methylation is inverses of the density of RNA polymerase II (Pol II), it was suggested that DNA methylation impedes transcriptional elongation by Pol II and predicted that DNA methylation plays a negative role in the regulation of transcription (Zilberman et al., 2007).

In animals, the majority of the studies of DNA methylation are carried out in mammalian systems, in which 60-90% of all the CpG sites are methylated (Ehrlich et al., 1982). Unlike in plants, most of the methylated cytosines in vertebrates are present within a context of CpG dinucleotides (Bird, 2002; Bird, 1986). Genome-wide mapping of DNA methylation indicated that some genomic regions, especially the regions that contains high enrichment of the CpG dinucleotides, prefer to be highly methylated, including repetitive elements (transposons, SINEs, LINEs), satellite DNA, intergenic regions and the exons of the genes. However, a conspicuous exception to the global distribution of DNA methylated is CpG islands (CGIs), which refer to a short un-methylated CpG rich region locates in the genome (Bird et al., 1985) (Figure 1.1B). Actually, it has been found that approximately 60% of the annotated genes contain the CGIs and the length of the CpG island ranges from hundreds of base pairs to several kilo-base pairs (Larsen et al., 1992; Saxonov et al., 2006; Zhu and Yao, 2009).

1.2.2 DNA methyltransferases

DNA methyltransferases (DNMTs) play a role in transferring and covalently attaching a methyl group onto the 5-position carbon of the cytosine pyrimidine ring by using *S*-adenosyl-*L*-methionine (SAM) as the methyl group donor. According to their function, there are two typical classes of DNA methyltransferases: one is involved in the establishment of DNA methylation, whereas the other is essential for the maintenance of DNA methylation (Chen and Riggs, 2011; Goll and Bestor, 2005). The first identified mammalian DNA methyltransferase was DNMT1, followed by DNMT2, DNMT3A, DNMT3B and DNMT3L (Bestor et al., 1988; Gruenbaum et al., 1982; Okano et al., 1998b; Van den Wyngaert et al., 1998; Yoder and Bestor, 1998).

All the DNMTs contain a conserved C-terminal domain, which is essential for

their methyltransferase activity, apart from DNMT3L (Cheng, 1995; Cheng and Blumenthal, 2008; Jeltsch, 2002; Jurkowska and Bal, 2001). The catalytic domain includes five conserved motifs, among which I and X are involved in the interactions between the DNMTs and their cofactors, whereas motifs IV, VI and VIII exert the "real catalytic" function. In addition, a non-conserved region, which is located in the middle of the catalytic domain, is required for accurate targeting (Cheng, 1995; Cheng and Blumenthal, 2008; Jeltsch, 2002). Furthermore, the N-terminal regions of DNMTs are suggested to be responsible for protein-DNA binding and protein-protein interaction (Jurkowska and Bal, 2001).

DNA methyltransferase 1 (DNMT1)

Bestor *et al* cloned and characterized the first mammalian DNMT from mouse (Bestor, 1988). Actually, the first identified DNA methyltransferase, DNMT1, is the most abundant DNA methyltransferase in mammalian cells. It has been found that the N-terminus of DNMT1 is essential for its stable expression and function *in vivo* (Tucker et al., 1996). In addition, since the N-terminus of DNMT1 contains several functional domains, including a proliferating cell nuclear antigen-binding domain (PBD), a nuclear localization signal (NLS), an cysteine-rich ATRX zinc finger DNA-binding motif and a Plant homeo domain (PHD), this domain is also involved in the discrimination of the hemimethylated from the unmethylated DNA strands.

During cell division, DNMT1 has been found to form a complex with the proliferating cell nuclear antigen (PCNA) and selectively localized to the replication foci in S phase (Iida et al., 2002; Leonhardt et al., 1992; Margot et al., 2001). Actually, apart from the centromere region, the chromatin loading patterns of DNMT1 spread all over the genome (Easwaran et al., 2004). Recently, it was found that the catalytic

activity of DNMT1 might be inhibited when its N-terminal region is bound to the unmethylated DNA (Bacolla et al., 1999). In addition, DNMT1 predominantly binds to the hemimethylated DNA and shows reduced specificity towards the unmethylated substrates (Fatemi et al., 2001). *In vitro,* DNMT1 displays 7- to 100-fold higher activity towards the hemimethylated DNA when compared to the methylated or the unmethylated double stranded DNA (Bestor, 1988; Fatemi et al., 2002; Gowher et al., 2005b; Pradhan et al., 1999).

Although DNMT1 has also been demonstrated to play a role in the establishment of DNA methylation by interaction with *de novo* DNA methyltransferase DNMT3A, it is considered to be the most important DNA methyltransferase for the maintenance of DNA methylation (Bestor, 1988; Fatemi et al., 2002; Gowher et al., 2005b). During mitosis, DNMT1 selectively binds to the parental strands with the help of UHRF1, a RING finger-associated protein, and establishes DNA methylation on the newly synthesized strands according to the patterns of the parental strands. Deletion of *Dnmt1* leads to the impairment of global DNA methylation in ES cells and embryos (Beard and Wilson, 1995; Danam et al., 2001; Li et al., 1993; Li et al., 1992). The *Dnmt1*$^{-/-}$ ES cells proliferate in culture, but undergo apoptosis upon differentiation (Li et al., 1992; Tsumura et al., 2006). In addition, deletion of *Dnmt1* is embryonic lethal and the *Dnmt1-/-* embryos die at embryonic day 9.5 (E9.5) (Beard and Wilson, 1995; Danam et al., 2001; Li et al., 1993; Li et al., 1992).

In mammals, there are three isoforms of DNMT1, including the somatic isoform of DNMT1 (DNMT1s), DNMT1b and an oocyte-specific isoform of DNMT1 (DNMT1o). Among these isoforms, DNMT1s is only present in the nucleus, whereas the DNMT1o is always synthesized and stored in the cytoplasm of mature oocytes.

However, after fertilization, the DNMT1o could be transferred into the nucleus and both maternal and zygotic DNMT1 are essential for the maintenance of genomic imprinting in embryos (Hirasawa et al., 2008). The expression levels of DNMT1 vary between different tissues and throughout the cell cycles *in vivo*. In general, DNMT1 is more abundant in the highly proliferating tissues and in S phase, there is relative more DNMT1 expressed in the cells (Robertson et al., 2000; Szyf et al., 2001). Additionally, DNMT1 is also associated with tumorigenesis and plays a role in both *de novo* and maintenance of DNA methylation in the cancer cells (Ting et al., 2006).

DNA methyltransferases 3A (DNMT3A) and 3B (DNMT3B)

Through a search of the expressed sequence tag (EST) databases with the conserved motifs of the cytosine DNA methyltranferases, three additional proteins were found, which share the homology within the catalytic domain of DNMT1, including DNMT2, DNMT3A and DNMT3B. Deletions of either *Dnmt3a* or *Dnmt3b* lead to the impairment of DNA methylation at repetitive elements and *de novo* methylation of the introduced retroviral DNA (Okano et al., 1999). In addition, as both the enzymes display activity towards the unmethylated DNA and play a role in the *de novo* DNA methylation during mammalian development (Okano et al., 1999), it has been suggested that both the DNMT3A and DNMT3B play an essential role in the establishment of the DNA methylation.

During cell divisions, since DNMT1 alone is unable to maintain all the methylated cytosines within the genome and *Dnmt3a* and *Dnmt3b* double knockout ES cells progressively lose DNA methylation at the repetitive elements, DNMT3A and DNMT3B are also suggested to play a role in the maintenance of DNA methylation (Liang et al., 2002). Furthermore, both enzymes have also been found to be able to

methylate the cytosines within non-CpG context (Gowher & Jeltsch, 2001; Ramsahoye et al., 2000). In mouse, there are three isoforms of DNMT3B. DNMT3B1 is highly expressed in ES cells, whereas DNMT3B2 and DNMT3B3 are expressed throughout most of the spermatogenesis (Oakes et al., 2007; Okano et al., 1998a).

Apart from the shorter N-terminal domain of DNMT3B, the amino acid sequences of DNMT3A and DNMT3B are quite similar and both of the two enzymes display no preference for methylating the unmethylated or the hemimethylated DNA *in vitro* (Okano et al., 1998a). The expression levels of these two enzymes are also quite similar and both of the two enzymes are highly expressed in early embryos and ES cells, but not highly expressed in somatic cells (Okano et al., 1998a). However, the knockout mice for either *Dnmt3a* or *Dnmt3b* display significant differences in phenotypes and physiological functions (Okano et al., 1999). For instance, knockout of *Dnmt3a* is postnatal lethal with the newly born mice dying between 4-8 weeks, whereas the *Dnmt3b-/-* mice die around embryonic day 14.5 (E14.5) and exhibit severe vascular and liver defects (Bachman et al., 2001; Hansen et al., 1999; Xu et al., 1999). Moreover, the functions of DNMT3A and DNMT3B in *de novo* DNA methylation are also different, since DNMT3A is essential for the establishment of germline imprinting, whereas DNMT3B plays a role in the establishment of DNA methylation at the minor satellite in mice and at the repetitive elements and pericentromeric heterochromatin regions in humans (Kaneda et al., 2004; Okano et al., 1999).

Similar to the *Dnmt1-/-* embryos, the embryos that lacking both the *Dnmt3a* and *Dnmt3b* display defects in the *de novo* DNA methylation during development and die around embryonic day 11.5 (E11.5) (Okano et al., 1999). However, the *Dnmt3a* and

Dnmt3b double knockout ES cells are viable, which indicated that DNA methylation is dispensable for the self-renewal of ES cells (Tsumura et al., 2006).

DNA methyltransferase 2 (DNMT2)

As mentioned above, DNMT2 is considered to be one of the DNA methyltransferases, since it contains a similar catalytic domain to DNMT1. However, DNMT2 displays extremely weak DNA methyltransferase activity (Hermann et al., 2003; Jurkowski et al., 2008; Tang et al., 2003). Deletion of *Dnmt2* had no discernable effect on the patterns of the global DNA methylation. Actually, DNMT2 displayed methyltransferase activity towards neither the unmethylated nor the hemimethylated DNA *in vitro*. A recent study found that DNMT2 methylates cytosine within the anticodon loop of the aspartic acid transfer RNA (tRNA-Asp), which indicates that DNMT2 plays a role as an RNA, rather than a DNA, methyltransferase in mammals (Ndlovu et al., 2011).

DNA methyltransferase 3L (DNMT3L)

DNMT3L is a catalyticaly inactive DNA methyltransferase. In mammals, it has been found that DNMT3L co-localizes with DNMT3A and DNMT3B *in vivo* (Nimura et al., 2006). In addition, DNMT3L could stimulate the activity of DNMT3A and DNMT3B (Chedin et al., 2002; Gowher et al., 2005a). Similar to DNMT3A, DNMT3L also plays a role in the establishment of genomic imprinting in the primordial germ cells (Hata et al., 2002; Bourc'his et al., 2001). Although both the male and female *DNMT3L* knockout mice are viable, its deletion results in sterility (Webster et al., 2005; Bourc'his and Bestor, 2001).

1.2.3 Functions of DNA methylation

The earliest literature, which suggested that DNA methylation could serve as an epigenetic mark in vertebrates, was published approximately 40 years ago (Holliday and Pugh, 1975; Riggs, 1975). In these papers, the authors proposed that the DNA sequences could be *de novo* methylated and there must be some enzymes, which play a role in the maintenance of DNA methylation. Furthermore, it was also suggested that the stable DNA methylation could be interpreted by DNA-binding protein and consequently, the presence of DNA methylation could directly result in transcriptional repression (Holliday and Pugh, 1975; Riggs, 1975). In evolutionary biology, DNA methylation is also thought to be essential for the acceleration of evolutionary processes, since spontaneous or catalytic deamination of methylated cytosine could create a thymine-guanine mismatch and the advantageous mismatch that occurred in the germ cells could be inherited by the next generation (Petersen-Mahrt and Neuberger, 2003).

The role of DNA methylation in gene regulation

In mammals, DNA methylation plays an essential role in the regulation of gene expression (Figure 1.1B), however the outcome of its role in the gene regulation is strongly associated with its positions. In general, DNA methylation that is present at the promoter of a given gene inhibits transcription, whereas the gene body DNA methylation plays an opposite role (Jones, 2012). In humans, there is no DNA methylation at the promoters of approximately 60% of the annotated protein coding genes. The unmethylated high CpG density regions that are present at these unmethylated promoters are known as CpG islands (CGIs) (Illingworth et al., 2010) (Figure 1A). It has been found that the genes that contain the CGIs are normally

characterized by the nucleosome-depleted regions (NDRs) around their transcription start site (TSS) and the NDRs are often flanked by the H3K4me3 marked nucleosomes (Kelly et al., 2010). However, genes that contain high levels of DNA methylated at their promoters also exist in humans, such as maternal imprinted genes and the genes that located on the inactive X chromosome in females (Figure 1B) (Jones, 2012). Currently, it is accepted that the high levels of DNA methylation at the promoters allow a long-term maintenance of the silenced state of the genes (Jones, 2012).

In contrast to the genes that contain the CGIs within their promoters, the CGIs-poor genes display variable patterns of DNA methylation at their promoters. For instances, there are low levels of DNA methylation at the promoters of some genes, which are highly transcribed in primordial germ cells (PGCs). In addition, some genes that selectively express in the ES cells contain high levels of DNA methylation around their promoters in the sperm, but not in oocytes or somatic cells (Farthing et al., 2008). However, the underlying functions of the DNA methylation that present at the promoters of the non-CGI genes are still obscure.

Being a feature of the transcribed genes, the early literature concluded that the gene body DNA methylation plays an active role for the transcription (Jabs et al., 1984). Recently, observations of the whole genome bisulfite DNA sequencing in plants and animals had confirmed the active correlations between the gene body DNA methylation and transcription (Cokus et al., 2008; Feng et al., 2010; Hellman and Chess, 2007; Lister and Ecker, 2009). In addition, it has been found that compared with the initiation of transcription, elongation is less sensitive to the gene body DNA methylation. Moreover, since actively transcribed genes also contain H3K36me3 at their gene bodies and H3K36me3 is able to recruit the DNMTs, the elongation process might even

contribute to the *de novo* DNA methylation in the gene bodies (Hahn et al., 2011). Furthermore, gene body DNA methylation also plays a role in silencing of the repetitive elements (Yoder et al., 1997).

The role of DNA methylation in genomic imprinting

Genomic imprinting is an epigenetic phenomenon, which refers to the mono-allelic expression of certain genes according to their parent-of-origin. In mammals, the majority of the imprinted genes are present in clusters and each cluster contains at least one non-coding RNA gene. In addition, all the imprinting clusters contain a unique differentially methylated region (DMR), which carries the parental allele-specific DNA methylation. Since loss of DNA methylation at ICRs causes aberrant expression of imprinted genes, these DMRs are known as the imprinting control regions (ICRs) and it has been found that the mono-allelic expression of imprinted genes is directly controlled by the DNA methylation that present at the ICRs (Shemer et al., 1997; Suzuki et al., 2002; Yatsuki et al., 2002) In addition, maternal and paternal ICRs located at different regions within the genome (Figure 1.2).

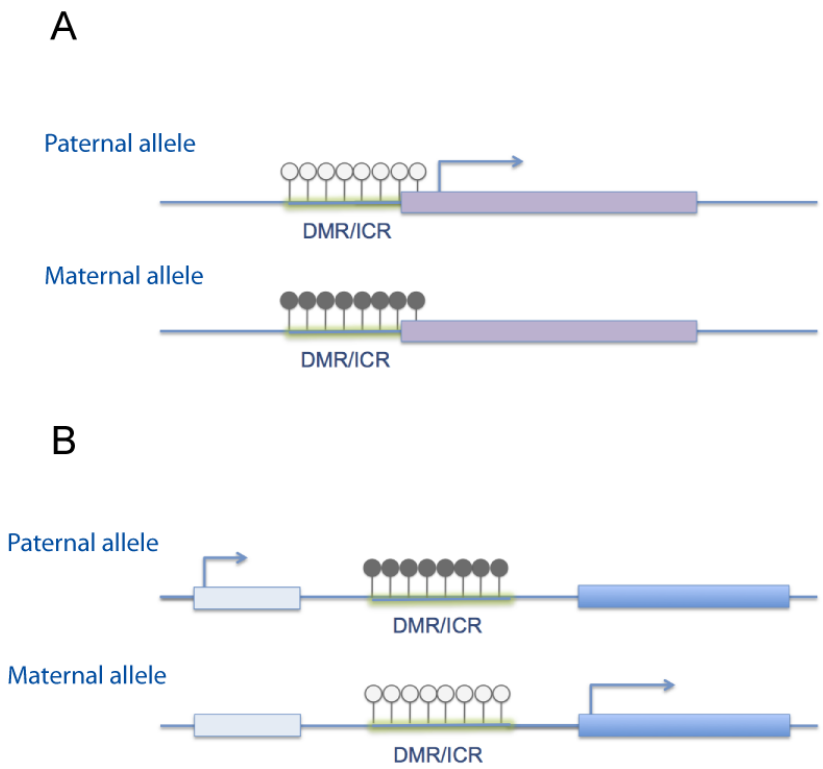


Figure 1.2 DNA methylation is essential for the maintenance of genomic imprinting.

In mammals, there are two types of imprinted DNA methylation present at ICRs. (A) A model of a maternally imprinted ICR. The maternally imprinted ICR is located at the promoters of the imprinted genes. For these loci, only the ICR presents on the maternally derived allele is methylated and consequently, the paternal copy of the maternally imprinted genes are expressed. In mouse, there are 16 maternal ICRs that have been identified. (B) A model of a paternally imprinted ICR. The paternally imprinted ICR is located in the intergenic regions of imprinted genes. Opposite to the maternal ICRs, only the ICR present on the paternally derived allele is methylated. The DNA methylation present at the paternally imprinted ICR regulates the imprinted genes that are located both upstream and downstream. In mouse, there are 3 paternally imprinted ICRs that have been identified. The white and black lollypops represent for the unmethylated and the methylated cytocines, respectively. The green shadow represents for the differentially methylated regions (DMR)/imprinted control regions (ICR). The arrow represents for the transcription of an imprinted gene.

Although the regulatory mechanisms for paternal and maternal ICRs in the regulation of the imprinted genes are different, mis-regulation of either paternal or maternal imprinted genes will lead to the severe disorders in humans (Cassidy and Driscoll, 2009). Therefore, it is essential to investigate how the imprinted DNA methylation is stably maintained in mammals. Previous studies have found that DNMTs are the main enzymes for the maintenance of the imprinted DNA methylation (Hirasawa et al., 2008). In addition, some other enzymes and DNA binding proteins also play a role in the maintenance of the imprinted DNA methylation (Li et al., 2008; Quenneville et al., 2011; Nakamura et al., 2012; Xin et al., 2003). However, take histone methyltransferase G9a for example, whether it plays a global role in the maintenance of imprinted DNA methylation and the underlying mechanism(s) for G9a in the maintenance of imprinted DNA methylation are still elusive.

The role of DNA methylation in X-chromosome inactivation

Placental mammals compensate the different dosages of X chromosome by randomly silencing one of the two X chromosomes in the female cells. To explain the mottled phenotype of female cats, which contain the heterozygous coat color genes on the X chromosome, the underlying mechanism of X chromosome inactivation was originally proposed in 1960s (Lyon et al., 1961). During the past 50 years, the mechanism of X chromosome inactivation has been the subject of numerous investigations. Today, it has been shown that the random X chromosome inactivation occurs at the late blastocyst stage and the majority of the genes located on the inactive X chromosome (Xi) are silenced (Ariel et al., 1995; Loh et al., 2007).

X chromosome inactivation is controlled by the expression of a noncoding RNA gene, called *Xist*. In adult females, *Xist* is one of the few genes that expressed from the

inactivated X chromosome (Xi). Since DNA methylation present at the X-pairing region (Xpr) plays a role in the regulation of *Xist*, it has been found that X chromosome inactivation is another epigenetic phenomena, which is directly controlled by DNA methylation (Loh et al., 2007). Recently, it has been found that the expression of *Xist* might also accumulate DNA methylation to the Xi and further condenses it. In addition, another noncoding RNA, named *Tsix*, has also been found to interact with *Xist* and both of these noncoding RNAs play a role in the X chromosome inactivation. In mammals, X chromosome inactivation occurs in the early developmental stage, however the timing of determination of X chromosome inactivation is different between species. In mouse, the X chromosome inactivation occurs in the late blastocyst stage (Xu et al., 2007).

The role of DNA methylation in genome stability

DNA methylation also plays a role in the maintenance of genome stability. The majority of the observations that indicate the DNA methylation functions in genome stability are carried out in human disorders. For instance, mutations within DNMT3B are found in patients with immunodeficiency, centromere instability and facial anomalies (ICF) syndrome and result in the reduction of the global levels of DNA methylation (Dodge at al., 2005; Ehrlich, 2003). Consistently, the hypomethylated cancer cells often display genome unstability. For instance, the mouse aggressive T-cell lymphomas contain extremely low levels of DNA methylation and are accompanied by a high frequency of trisomy of chromosome 15 (Eden et al., 2003).

The possible explanation for the role of DNA methylation in genome stability could be the DNA methylation is essential for the accurate segregation of the chromosomes. Therefore, loss of DNA methylation might lead to the initiation or progression of the tumorigenesis. Additionally, since DNA methylation also plays a role

in the suppression of retrotransposons, the role of DNA methylation in genome stability might also be achieved through preventing the transposition of retrotransposons and protecting the chromosomes from aberrant recombination.

1.2.4 Dynamics of DNA methylation

Compared with histone modifications, DNA methylation is a relatively stable epigenetic mark. However, during the early embryonic development and the maturation process of germ cells, DNA methylation undergoes dynamic reprogramming (Figure 1.3). In general, there are two waves of global DNA demethylation occurring in the developing embryos, with the first one commences soon after the fertilization and the second one occurs in the primordial germ cells (PGCs) in the embryos at E13.5 (Feng et al., 2010).

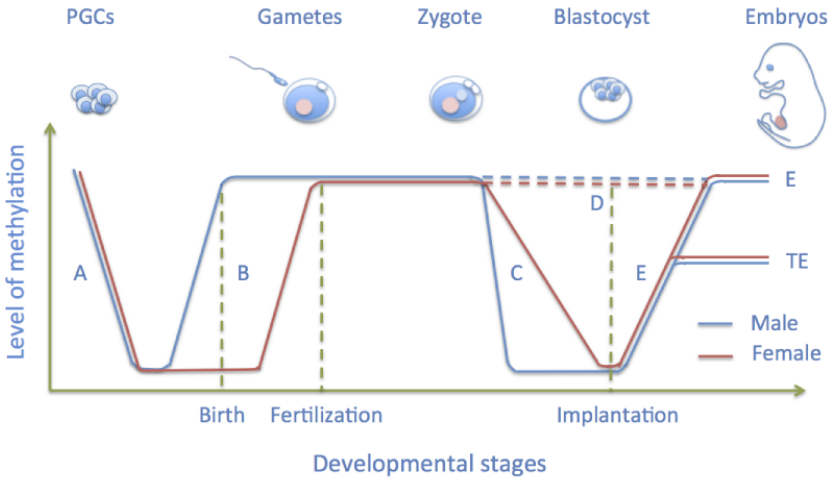


Figure 1.3 Dynamic of DNA methylation during gametogenesis and early embryonic development.

(A) During the early stage of primordial germ cells (PGCs) development, genome-wide DNA demethylation occurs, including the DNA methylation that present at the imprinted control regions (ICRs). By the time of the PGCs enter the genital ridge, only the DNA methylation that presents at the repetitive element is maintained. (B) During gametogenesis, global DNA methylation is re-established, including the new sex-specific imprinted DNA methylation present on the differentially methylated regions (DMRs) of the maternal and paternal chromosomes. However, the re-establishment of imprinted DNA methylation in the male and female germ cells occurs at different times of development. In male germ cells, the paternal imprinted DNA methylation is initially established in prospermatogonia during the period between mitotic arrest and birth. By contrast, the maternal imprinted DNA methylation is established in the growing oocytes after birth and completed before ovulation. (C) In the period between fertilization and the blastocyst stages, the global DNA methylation undergoes another wave of DNA demethylation. In this process, the DNA methylation in the paternal and maternal pro-nuclei undergoes active and passive DNA demethylation respectively, since the loss of paternally derived DNA methylation is independent of cleavage. (D) The imprinted DNA methylation established in the germline is stably maintained during the second wave of DNA demethylation (dashed lines). Actually, the normal development of the embryos requires the correct reading of the genomic imprinting. (E) Around the time of implantation, re-establishment of the global DNA methylation takes place and the patterns of DNA methylation are faithfully inherited by the daughter cells throughout the development of the individual. E: embryo; TE: trophectoderm. Figure is adapted from Ishida and Moore, 2012.

During the stage from zygote to blastocyst, both paternally and maternally derived chromosomes undergo DNA demethylation. Since paternal genome undergoes a rapid and cell division-independent DNA demethylation, it has been proven that DNA methylation presents on the paternally derived genome is erased through an active DNA demethylation. By contrast, maternally derived genome is thought to undergo a passive DNA demethylation, since its DNA methylation undergoes a cell-division dependent DNA demethylation. Actually, after fertilization, the global level of DNA methylation in the maternally derived genome is continuously being reduced during cleavages until the late blastocyst stage. Interestingly, a recent study indicated that the oocyte genome displays substantial hypomethylation even before fertilization, which raises the question of when the DNA demethylation of the maternally derived genome actually occurred (Smith et al., 2012).

However, some specific genomic loci, such as the germline differentially methylated regions (DMRs), have been found to be resistant to the first wave of DNA demethylation that occurs soon after fertilization. It has been shown that compared with the DNA methylation that presents at the boundary regions of DMRs, imprinted DNA methylation that established in the germ cells is maintained or at least partially maintained during the first wave of DNA demethylation (Kobayashi et al., 2012; Smith et al., 2012). In addition, a proportion of the DNA methylation that presents at the promoters of oocyte-specific genes is also protected, which suggested that germline DMRs are not the only regions that escape from DNA demethylation in the early embryos (Venter et al., 2011).

Since the DNA methylation present at the repetitive elements is essential for the maintenance of genome stability, for a long time, it was believed that this DNA

methylation should be protected during the first wave of DNA demethylation. However, recent observations indicate that both the long interspersed repeats (LINEs) and the short interspersed repeats (SINEs) partial lose DNA methylation by the blastocyst stage (Kobayashi et al., 2012; Smith et al., 2012). Admittedly, these observations imply that the DNA methylation present at the different genomic regions might be distinctly regulated during the early mammalian development.

The first wave of DNA demethylation terminates at the blastocyst stage and the global DNA methylation by this point reaches the lowest level. Upon implantation, *de novo* DNA methylation occurs in the epiblast with the silencing of the majority of the germline-expressed genes (Isagawa et al., 2011). In addition, by gaining of the DNA methylation at promoters, some tissue specific genes in the pluripotent inner cell mass (ICM) are also silenced during this stage. Together with the DNA methylation that is present on the inactive X chromosome, DNA methylation established during this stage plays an essential role in the further embryonic development (Isagawa et al., 2011).

The second wave of DNA demethylation occurs in the primordial germ cells (PGCs) followed by a wave of *de novo* DNA methylation. In mice, PGCs arise from approximately 45 cells in the epiblast at around embryonic day 7.5 (E7.5). PGCs then migrate through the hindgut and extra-embryonic regions and finally arrive at the primordial gonad in the proximity of the genital ridge by E10.5 to E11.5 (Buehr, 1997). It has been found that nearly the entire genome loses DNA methylation in the developing PGCs, including the germline DMRs. After removal of the parentally inherited DNA methylation, a wave of *de novo* DNA methylation occurs globally and consequently, DNA methylation is established at the appropriate regions within the genome (Kagiwada et al., 2013). These also include the germline imprinting control

regions (ICRs), where the allele-specific patterns of DNA methylation are deposited depending on the sex of the developing germ cells (Kagiwada et al., 2013). Notably, once established, the DNA methylation that presents at the germline DMRs is stably maintained throughout life.

1.3 Genomic imprinting

The diploid organisms carry two copies of the genome, with one copy inherited from each parent. In mammals, for the vast majority of genes, the expression occurs from both copies. However, there is a small group of genes whose expression only occurs from one copy of the two genes and these genes are called imprinted genes. However, recent studies indicated that nearly 10% of the autosomal genes can also be mono-allelically expressed in mammals and these randomly monoallelically expressed genes can lead to unique cell identity (Chess, 2012).

In mammals, nearly 80% of the imprinted genes are located in the imprinting clusters and each imprinting cluster contains a germline imprinting control region (ICR). Since removal of parent-specific DNA methylation from the germline ICRs leads to aberrant expression of imprinted genes, it has been found that mono-allelically expression of imprinted genes is controlled by the DNA methylation that presents at the germline ICRs. However, apart from the DNA methylation that presents at the germline ICRs, the DNA methylation that presents at the secondary DMRs is also involved in the regulation of the imprinted genes.

1.3.1 Discovery of imprinted genes

In the 1980s, by carried out the pronuclear transplantation experiments, it has been found that maternally and paternally uniparental embryos can only develop into

the predominantly embryonic tissues or larger extra-embryonic tissues, respectively (Surani and Barton, 1983). These observations indicate that normal mammalian embryonic development requires both maternally and paternally derived genome. Furthermore, it has also been found that accurate expression of imprinted genes, which locate on the maternally and paternally derived chromosomes, actually plays an essential role in the normal development of the embryos. However, a recent study demonstrated that a mature oocyte combined with a maternally derived nucleus could also generate viable mice (Kawahara et al., 2007). In addition, using the non-growing oocytes, which contain mutations at two individual imprinted loci as the nuclei donor, it has been found that the frequency of the viable offspring is associated with the dosage of the imprinted genes (Kawahara et al., 2007). Taken together, these observations suggest that in mammals, the accurate expression of the imprinted genes is essential for the normal embryonic development.

The first identified imprinted gene, *Igf2r*, encodes the insulin-like growth factor type 2 receptor (Barlow et al., 1991). Apart from *Igf2r*, other imprinted genes were also identified in gene targeting experiments, such as *Igf2*, the insulin-like growth factor type 2 (DeChiara et al., 1991) and *Mest*, also known as *Peg1* (Lefebvre et al., 1998). Based on a different strategy, maternal imprinted gene *H19* was identified using single nucleotide polymorphisms (SNPs) that reside in the *H19* RNA (Bartolomei et al., 1991). Today, since a great number of the SNPs have been found in various mouse strains, the whole transcriptome analysis has been employed to identify the potential imprinted genes. This led to the identification of many monoallelically expressed RNAs (Babak et al., 2008; Wang et al., 2008). Currently, there are at least 80 imprinted genes known in humans and mice. Notably, the vast majority of the imprinted genes are distributed in

clusters and the average length of each cluster is about 1Mb. In addition, nearly every cluster contains paternal and maternal imprinted genes and at least one gene, which encodes for a non-coding RNA.

1.3.2 Functions of imprinted genes

In mammals, imprinted genes are predominantly expressed in the prenatal stage and down regulated after birth. Within the embryos, imprinted genes are expressed in the placenta and brain, which is consistent with their functions in controlling of the embryonic growth, the development of the nervous system and behaviour (Davies et al., 2008). The main function of imprinted genes is to maintain the normal development of the embryos and the placenta. It has been found that silencing of *Igf2* causes aberrant transport of nutrients from the placenta to the growing embryo (Sibley et al., 2004). In addition, paternal imprinted gene *Ascl2* is responsible for the normal differentiation of the spongiotrophoblast cells in the developing placenta (Guillemot et al., 1995). Furthermore, paternal imprinted genes *peg10* and *Rtl1* are required for the normal development of the placenta (Ono et al., 2006; Sekita et al., 2008).

Consistently, aberrant expression of imprinted genes leads to the abnormal embryonic growth. It has been found that the mis-regulation of *Delta-like homolog 1* (*Dlk1*) in embryos causes over-growth of the fetus, although there is no change of its expression in the placenta (da Rocha et al., 2009). *Dlk1* also plays a role in the lineage specification, since mutations of *Dlk1* in mice lead to growth retardation, skeletal abnormalities, adipose tissue defects and aberrations in the hematopoietic lineage cells (Raghunandan et al., 2008). Another imprinted gene, Cyclin dependent kinase inhibitor 1c (*Cdkn1c*), is intensively expressed in the somatic tissues during development, including muscle, lung, kidney and eye in mammals. Since CDKN1C plays an essential

role in the regulation of cell cycle and knockout of *Cdkn1c* in mice leads to placentomes, spongioblastoma, proliferation of labyrinthine and fibrin deposition of the intervillous space, it has been found that *Cdkn1c* plays an essential role in the early stages of mammalian development (Takahashi et al., 2000).

Apart from playing a role in the development of the placenta and the embryos, imprinted genes are also required for the normal development of the brain. It has been found that mutations of imprinted genes lead to aberrant behaviour and mental retardation in mammals. In mice, apart from growth retardation, females that lack either *Peg1* or *Peg3* display abnormal maternal behaviour (Lefebvre et al., 1998). In addition, imprinted genes have also been found to play a role in energy homeostasis. One of the best instances is paternal imprinted gene $G_s\alpha$, which encodes a G protein that controls the production of cAMP. It has been found that the abnormal expression of $G_s\alpha$ from the maternal allele in the brain leads to the glucose intolerance, insulin resistance and obesity in mice (Weinstein et al., 2010).

1.3.3 Regulation of mono-allelic expression of imprinted genes

As discussed above, the maternal imprinted genes always contain a differentially methylated region (DMR) at their promoters. Without any exception, these DMRs are unmethylated on the active allele, but methylated on the inactive one. By contrast, there is no DMR that have been found at the promoters of paternal imprinted genes. However, paternal imprinted clusters contain an intergenic germline DMR. The distribution of these DMRs around the imprinted genes raised the hypothesis that mono-allelically expression of imprinted genes might be regulated by the DNA methylation that present at the relevant DMRs. In mice, targeted deletion studies suggest that the mono-allelically expression of the imprinted genes is directly controlled by the imprinted

DNA methylation (Wutz et al., 1997) and all the germline DMRs are likely to regulate the imprinted genes in a *cis*-regulatory pathway. However, the paternal and maternal imprinted genes are regulated by the different mechanisms.

The best studied paternal imprinted genes located at the *Igf2-H19* locus. Although *Igf2* and *H19* are located 90 kb away in the genome, they share the same enhancer, which locates at downstream of *H19*. Normally, *Igf2* is expressed from the paternal allele, whereas *H19* is expressed from the maternal allele (DeChiara et al., 1991). It has been found that a germline DMR that resides 2-4 kb upstream from the TSS of *H19* plays a direct role in the regulation of the monoallelic expression of these two genes (Tremblay et al., 1997). In addition, this germline DMR was found to contain multiple binding sites for a zinc finger insulator protein CTCF and CTCF only binds to the unmethylated germline DMR, which presents at the *Igf2-H19* locus on the maternal allele (Bell and Felsenfeld, 2000). Since the binding of CTCF to the unmethylated germline DMR inhibits the interaction between enhancer and promoter of *Igf2*, maternal copy of *Igf2* is inactive. By contrast, the enhancer induces expression of *H19* on the maternal allele. Without binding of CTCF to the methylated DMR on the paternal allele, enhancer induces expression of *Igf2*, whereas *H19* is not expressed in this situation.

Since maternal germline DMRs are located at promoters, compared to the paternal imprinted genes, it has been found that the maternal imprinted genes are regulated by a different mechanism. For instance, germline DMR located within an intron of *Igf2r* overlaps with the promoter of *Airn*, which encodes a large non-coding RNA and is exclusively expressed from the paternal allele (Wutz et al., 1997). It has been found that the expression of *Airn* correlates with the methylation state of its promoter (Sleutel et al., 2002). In addition, the non-coding RNA has also been found to play a role in the

regulation of the imprinted genes located at the *Kcnq1*, *Snrpn* and *Gnas* clusters (Chotalia et al., 2009).

1.3.4 Establishment of germline imprinting

Among the approximately 80 identified imprinted genes that reside in 26 different imprinting clusters, only 3 paternal imprinting clusters are found: *Igf2-H19*, *Dlk1-Dio3* and *Rasgrf1* loci. Although maternal and paternal germline DMRs are located at different genomic regions, all the imprinted DNA methylation is established during gametes development (Figure 1.4). In male mammals, the establishment of paternal imprinted DNA methylation occurs during the period of mitotic arrest and birth in the diploid gonocytes, whereas in female mammals, the maternal imprinted DNA methylation is established in the developing oocytes that are arrested in the diplotene stage of meiosis prior to ovulation (Davis et al., 1999; Lucifero et al., 2002).

Apart from the *Rasgrf1* germline DMR, it has been found that all the imprinted DNA methylation that present at the ICRs is established by *de novo* DNA methyltransferases DNMT3A and DNMT3L (Bourc'his and Bestor, 2006). However, how the germline DMRs are accurately recognized and discriminated from the rest of the genome is still poorly understood. By X-ray crystallography, it has been found that DNMT3A and DNMT3L prefer to form a tetrameric complex, which contains of two copies of each enzyme and this complex specifically targets to a pair of CpG dinucleotides, which are separately 8-10 bp (Smith et al., 2007). In the maternal imprinted DMRs, this specific spacing has already been identified, although it is also comprehensively distributed across the whole genome. In addition, DNMT3L may play a role in chromatin loading of the DNMT3A/DNMT3L complex at the imprinted DMRs, since DNMT3L only interacts with the histone H3 that lacks of the H3K4 methylation

(Ooi et al., 2007). Furthermore, transcription through the imprinted DMRs in the oocytes has also been found to be essential for the accurate establishment of DNA methylation that presents at the maternal imprinted loci (Chotalia et al., 2009).

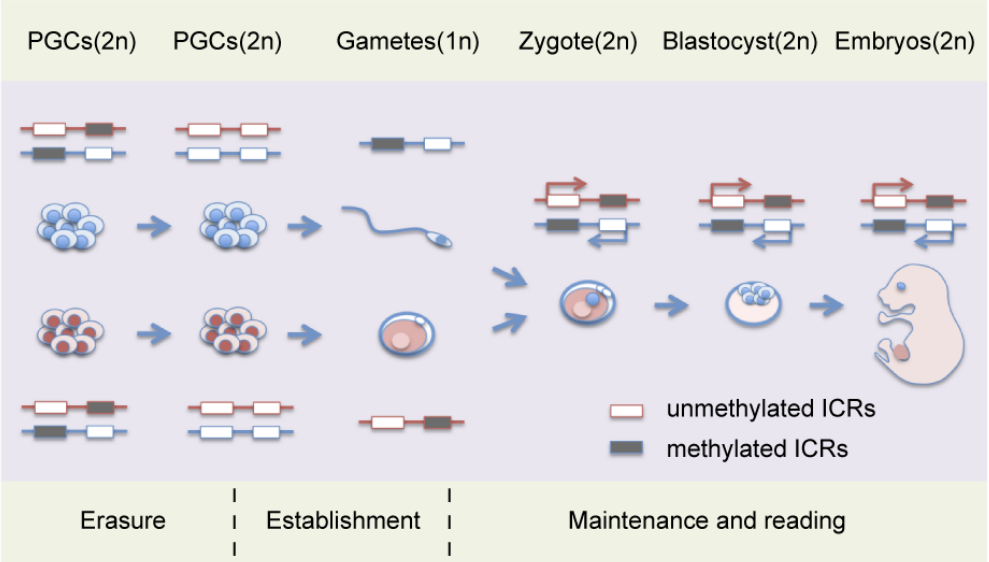


Figure 1.4 The establishment and maintenance of genomic imprinting.

In mammals, the gametes are derived from the primordial germ cells (PGCs). Initially, the PGCs contain the normal status of imprinted DNA methylation. At the very beginning stage of the gametogenesis, nearly all the DNA methylation in the PGCs is erased, including the imprinted DNA methylation. Next, during the maturation process of the PGCs, a wave of global *de novo* DNA methylation occured and consequently, both the mature oocytes and sperm contain the sex-specific pattern of imprinted DNA methylation. After fertilization, although another wave of DNA demethylation occurs, the imprinted DNA methylation is stably maintained in the preimplantation embryos. Figure is adapted from Rakyan et al, 2001.

1.3.5 Maintenance of genomic imprinting

Once the imprinted DNA methylation is established in the germ cells, the patterns of imprinted DNA methylation are stably maintained throughout life (Figure 1.4). Since inappropriate expression of the imprinted genes leads to the defects of the embryonic development and causes the human disorders, it is important to investigate how the imprinted DNA methylation is maintained. The most challenging period for the maintenance of genomic imprinting is in the early developmental stage, since both paternally and maternally derived chromosomes undergoes DNA demethylation. Strikingly, although most of the global DNA methylation is erased and subsequently reestablished, the DNA methylation that presents at the germline DMRs is stably maintained during this period. In addition, the patterns of imprinted DNA methylation are also inherited precisely by the newly synthesized genome during the cleavages. These observations indicate that there must be some specific mechanisms, by which imprinted and non-imprinted DNA methylation could be discriminated and only the imprinted DNA methylation is maintained in the pre-implantation embryos.

Currently, several enzymes have been found to play a role in the maintenance of imprinted DNA methylation. DNMT1 is the first identified factor, which is required for the maintenance of imprinted DNA methylation. In the pre-implantation embryos, apart from the existence of a somatic form of DNMT1 (DNMT1s), an oocyte specific form of DNMT1 (DNMT1o) is also present and both of DNMT1o and DNMT1s are suggested to play a role in the maintenance of imprinted DNA methylation. However, a recent study indicated that DNMT1s alone is sufficient for the maintenance of imprinted DNA methylation in the early embryos (Howell et al., 2001; Hirasawa et al., 2008). In addition to DNMT1, DNMT3A and DNMT3B are also been found to play a role in the

maintenance of genomic imprinting, since deletion of *Dnmt3a* and *Dnmt3b* in ES cells leads to the DNA demethylation at the imprinted loci (Chen et al., 2003; Liang et al., 2002; Okano et al., 1999).

Apart from DNMTs, a KRAB zinc finger protein ZFP57 has also been found to play a role in the maintenance of imprinted DNA methylation (Temple and Shield, 2002). The role of ZFP57 in the maintenance of imprinted DNA methylation was initially discovered in patients with transient neonatal diabetes (TND) (Mackay et al., 2008). In addition, it was found that deletion of *Zfp57* even leads to the embryonic lethality and causes the mis-regulation of the imprinted genes (Li et al., 2008; Quenneville et al., 2011). In mouse ES cells, ZFP57 and its cofactor KAP1 exclusively bind to the methylated allele and recruit SETDB1 to selectively establish H3K9me3 on the methylated allele at imprinted loci (Quenneville et al., 2011). Furthermore, it was shown that ZFP57 maintains the imprinted DNA methylation through recognition of a methylated hexanucleotide sequence, TGCCGC, which presents at some of the ICRs (Quenneville et al., 2011).

In addition, STELLA, also known as PGC7 or GADD3, has also been found to play a role in the maintenance of genomic imprinting (Nakamura et al., 2007). STELLA was first found to protect the imprinted DNA methylation in embryos. Although the *Stella-/-* oocytes contain the normal patterns of imprinted DNA methylation, embryos derived from these oocytes display hypomethylation at multiple imprinted DMRs (Nakamura et al., 2007). Recently, it was shown that STELLA binds to the H3K9me2, which presents at the maternally derived genome and germline DMRs within the paternally derived chromosomes, and protects the DNA methylation at ICRs from TET3-mediated active DNA demethylation in the mouse zygote (Nakamura et al., 2012).

Furthermore, the members in the CATERPILLER protein family have also been found to play a role in the maintenance of imprinted DNA methylation. Female patients, who suffer the familial recurrent hydatidiform mole, a rare maternal effect autosomal recessive disease, display the abnormal patterns of DNA methylation at some of the germline DMRs in their oocytes. Since mutations in NLRP7 protein lead to the hydatidiform mole, it is likely that NLRP7 plays a role in the maintenance of genomic imprinting (Kou et al., 2008). Moreover, investigation of two siblings with Beckwith-Wiedemann syndrome implicated that NLRP2, another member in the CATERPILLER protein family, plays a role in the maintenance of imprinted DNA methylation at the *KCNQ1* cluster (Moser et al., 2009).

Recently, two Rb-binding protein-related genes, *Arid4a* and *Arid4b*, were also found to be involved in the maintenance of imprinted DNA methylation at the Prader-Willi syndrome imprinting center (PWS-IC) and Angelman syndrome imprinting center (AS-IC) (Wu et al., 2006). Interestingly, mice combine the homozygous deficiency of *Arid4a* and the heterozygous deficiency of *Arid4b* display reduced levels of H4K20me3 and H3K9me3 at the PWS-IC and AS-IC. Additionally, since maintenance of DNA methylation at the PWS-IC in ES cells requires the presence of G9a and one of the imprinted genes locate in the PWS-IC, named *Snrpn*, is bi-allelically expressed in the *G9a*-/- ES cells (Xin et al., 2003), it is suggested that G9a might also play a role in the maintenance of imprinted DNA methylation in mammals.

1.3.6 Genomic imprinting and human disorders

Since epigenetic mechanisms play an essential role in many biological events, it is undoubted that inappropriate alteration of epigenetic modifications might lead to the human disorders. Actually, after the mechanisms of the Prader-Willi syndrome and the

Angelman syndrome have been revealed, several human genetic diseases have been reported, which are associated with the mutations of imprinted genes. Although these disorders include various types of disruptions, such as the chromosomal duplication and deletion, most of them are caused by the aberrant dosage of the imprinted genes.

Prader-Willi syndrome (PWS)

PWS is a multi-manifestation human genetic disorder. The patients of the PWS display variant pathological characteristics, including abnormal development in infancy, mental retardation, aberrant growth hormone secretion, hypogonadism and even obesity (Cassidy and Driscoll, 2009). As the first discovered human imprinting disorder, the estimated frequency of the PWS is about one in 15,000 to 25,000 (Buiting, 2010). The phenotype of PWS patients includes mild intellectual disability, lower birth weight, poor suckling and hypotonia to the voracious appetite (Cassidy and Driscoll, 2009). Some of the PWS patients contain an imprinted defect, which is associated with the failure to erase of the grandparental derived DNA methylation from their paternal germ cells (Buiting, 2010). Although 70% of the PWS have been reported to have a loss of paternally derived chromosome region 15q11-13 and 25% of them display maternal uniparental disomy (UPD) of chromosome 15, it is still unclear which specific gene(s) contribute to the phenotype of PWS (Buiting, 2010). Recently, a gene that codes a small nuclear RNA (snoRNA), named *SNORD116*, was suggested to play a role in the aetiology of PWS (Ding et al., 2008).

Angelman syndrome (AS)

The clinical phenotypes of the AS in humans involve severe intellectual disability, delayed weaning, microcephaly, frequent smiling and laughing (Buiting,

2010). Similar to the PWS, 70% of the AS patients display deletions in the proximity of the 15q11-13 locus. Interestingly, absence of 15q11-13 locus in majority of the AS patients occurs on the maternal allele. In addition, 10% of the patients contain a mutation within the maternal imprinted gene, *UBE3A*, which encodes an E3 ubiquitin ligase (Buiting, 2010). Furthermore, 2-4% of the AS patients display the abnormal patterns of the imprinted DNA methylation within the 15q11-13 locus and a small proportion of the patients carry a bi-directional imprinted control region (ICR) (Horsthemke and Wagstaff, 2008).

Silver-Russell syndrome (SRS)

SRS occurs in approximately one in 3,000 to 100,000 children and the patients display characteristic pre- and/or postnatal growth restriction, skeletal asymmetry and small triangular shaped face (Abu-Amero et al., 2008). Mutations within the 7q32 and 7p12.2-3 chromosomal regions that contain imprinted genes *MEST* and *GRB10*, respectively, are responsible for the SRS (Abu-Amero et al., 2008). In addition, since the SRS patients also display abnormal patterns of imprinted DNA methylation at other ICRs, it has been found that the SRS patients are unable to establish and maintain the overall patterns of imprinted DNA methylation properly (Kannenberg et al., 2012).

Beckwith-Wiedemann syndrome (BWS)

Down-regulation of the imprinted genes that locate at the 11p15.5 locus is involved in 85% of the BWS and the incidence of BWS is estimated about one in 13,700 in the live births (Gicquel et al., 2005). Opposite to the SRS, the phenotypes of the BWS patients include pre- and/or postnatal overgrowth, placental overgrowth, macroglossia and embryonal carcinoma (Choufani et al., 2010). In addition, around half

of the BWS patients show hypomethylation at the *KvDMR1* locus, whereas 5% of the patients display hypermethylation at the *H19-Igf2* ICR (Miho et al., 2012). Although most of the BWS cases are sporadic, 15% of the patients have inherited a mutation, among whom, 40% contain the mutations within the *CDKN1C* locus (Choufani et al., 2010).

In addition to the ICRs, mutations of the epigenetic marks that occurred at the other genomic regions can also lead to the human disorders. It has been found that alterations of DNA methylation and histone acetylation that present at the gene body regions could cause human prostate cancer (Cang et al., 2009; Kloth et al., 2012). In addition, other studies have found that a gene called mixed lineage leukemia (*Mll)* could cause the AML in humans by rearranging and fusing with the other genes that present on the different chromosomes. Furthermore, loss of DNA methylation that presents at the polymorphic D4Z4 repeat region on the chromosome 4 could also cause the facioscapulohumeral muscular dystrophy (FSHD) (van Overveld et al., 2003). Moreover, aberrant DNA methylation that presents at the retrotransposons could also lead to the human disorders (Sekigawa et al., 2003). Taken together, it is important to investigate the underlying mechanisms that link the epigenetic marks with the human disorders.

1.4 Histone modifications

In eukaryotes, the nucleosome serves as the basic component for the packaging of DNA into the chromatin. The core of a nucleosome is an octameric protein complex, which contains two copies of each histone, including H2A, H2B, H3 and H4. Since the tails of histones can further be modified (Figure 1.5), the structure of chromatin is strongly associated with the local histone modifications and varies in the nucleus. In

general, there are two categories of histone posttranslational modifications (HPTMs). The first includes the small chemical groups, such as the acetyl, methyl and phosphate groups. By contrast, the other involves the larger peptides, including the ubiquitin and the SUMO.

Histone post-translational modifications were first identified in the 1960s. However, the underlying functions of these marks were poorly understood for a long time. In recent years, many studies have shown that similar to DNA methylation, the HPTMs also play an essential role in the regulation of gene expression. Today, two underlying mechanisms of the HPTMs are considered. First, the HPTMs *per se* may facilitate the direct interaction between the effector proteins and the chromatin. Second, the HPTMs may alter the structure and block the accessibility of the chromatin at a given locus. Till now, a great number of the HPTMs have been found in mammals (Figure 1.5), among which some are transient, whereas the others are likely to be epigenetically heritable.

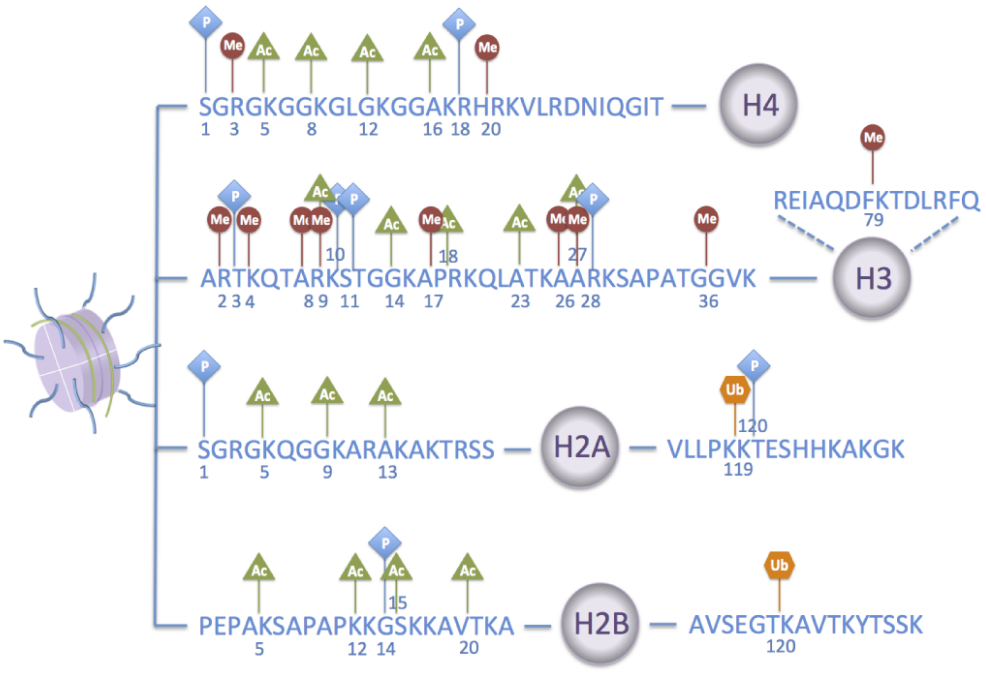


Figure 1.5 The schematic view of histone modifications.

The histone octamer contains two copies of each of the four histones, including H2A, H2B, H3 and H4. The tails of these histones can be modified *in vivo*. Although the majority of the modified residues contain one specific mark, a few of them can be modified differentially. Take lysine 9 and 27 on histone H3 for example, these two residues can either be methylated or acetylated. In general, both the acetylation and phosphorylation play an active role for gene expression, whereas the methylation and ubiquitination can either activate or repress the expression of the associated genes. Me: methylation; Ac: acetylation; P: phosphorylation; Ub: ubiquitination. Figure is adapted from http://www.epitomics.com (histone modificatios).

1.4.1 Distribution of histone modifications

Similar to DNA methylation, the distribution of HPTMs also varies across the chromatin and active and silenced genes contain different profiles of HPTMs (Figure 1.6). For a given gene, HPTMs are not only present at promoter or enhancer, but also in the gene body region (Figure 1.6) (Barth and Imhof, 2010). For instance, H3K4me2 and H3K4me3 usually exist at the promoters of actively transcribed genes and H3K4me1 and H3K27Ac normally co-exist at the enhancers (Sim III and Reinberg, 2006; Zentner et al, 2011). By contrast, H3K9me2 and H3K9me3 are usually present at the inactive genes and the peripheral regions within the nucleus (Barski et al., 2007). Furthermore, H3K27me3 and H4K20me3 are also normally present at the inactive genes (Cao et al., 2002; Rice et al., 2003). According to the distinct profiles of the HPTMs, it has been suggested that the positioning of the HPTMs is strongly associated with their functions.

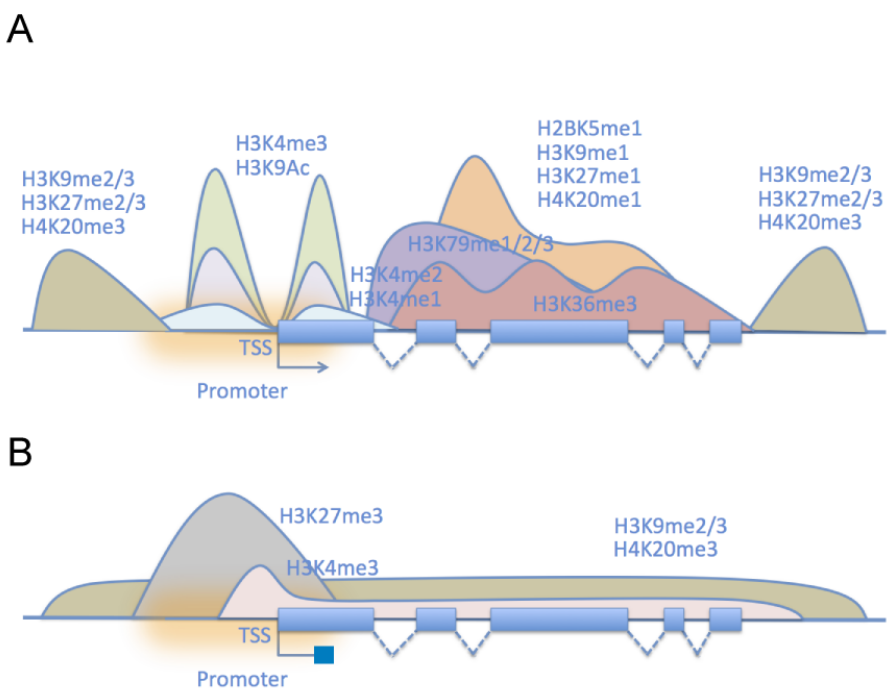


Figure 1.6 The schematic view of the distribution of histone modifications on active and repressive genes.

In mammals, the actively transcribed and transcriptional repressed genes contain distinct histone modifications at both the promoters and gene body regions. (A) In general, the promoters of active genes contains H3K4 methylation and H3K9 acetylation. At the gene body region, the presence of H3K36me3 serves for efficient elongation. Apart from H3K36me3, some other methylation marks, such as H4K20 and H3K79 methylation, also exist at the gene body region. Since most of these marks only contain one methyl group, it is believed that the mono-methylation seems to play an active role in transcription. (B) By contrast, the inactive gene normally contains H3K9me2/3, H4K20me3 and H3K27me3. However, the distribution of these repressive marks are discernable. For instances, H3K9me2/3 and H4K20me3 evenly exist through the gene region, whereas H3K27me3 is extremely enriched at the promoters. Note, H3K4me3 only exists at the promoters of bivalent genes. The yellow shadow represents for the promoter and TSS stands for the transcription start site. Figures are adapted from published article, Teresa K. Barth and Axel Imhof, 2010.

However, some observations demonstrate that the distribution profiles of the HPTMs are inconsistent in some cases. For example, together with the heterochromatin protein 1 gamma (HP1gamma), the repressive marks H3K9me2 and H3K9me3 have also been found at some active genes and the positioning of both the HP1gamma and the H3K9me2/3 depends on the transcription elongation (Vakoc et al., 2005; Vakoc et al., 2006). Additionally, some other repressive marks are also present at the active genes, including H3K27me1 and H3K20me1 (Carrozza et al., 2005; Vakoc et al., 2006). Furthermore, it has been found that H3K36me3 is enriched in the gene body regions and facilitates the transcription elongation (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005; Kizer et al., 2005). Another interesting observation is the bivalent chromatin state, where both the H3K4 and H3K27 methylation co-exist at the promoters of many protein coding genes in human and mouse ES cells (Cankovic et al., 2007). Since the bivalent domains tend to coincide with the transcription factor (TF) associated genes, the possible mechanism is that the bivalent domains suppress the differentiation associated genes in the pluripotent cells, while keeping them poised for activation in the differentiated cells (Bernstein et al., 2006).

1.4.2 Histone modifiers and the functions of histone modifications

In eukaryotes, histone modifications were first identified approximately 50 years ago (Murray, 1964) and actually, histone was the first protein that had been revealed to contain the ubiquitin mark. Although the correlations between the histone modifications and the transcriptional activity have been observed, the functions of these modifications were still obscure. Today, it has been found that histone modifications could regulate the gene activity in either direct or indirect ways. In addition, some histone modifications are only associated with gene expression, such as acetylation and

phosphorylation, whereas the others can play either a positive or a negative role in the regulation of the gene expression.

Histone lysine methyltransferases (KMTs), demethylases (KDMs) and the functions of histone lysine methylation

Although histone methylation was first discovered in the 1960s (Murray, 1964), histone lysine methyltransferases (KMTs) and histone lysine demethylases (KDMs) have only been identified relatively recently. The first HKMTs, SUV39H1 and its fission yeast homolog Clr4, were identified in 2000 (Rea et al., 2000), and since then a great number of enzymes have been found to exert a KMT function. All of these enzymes use *S*-adenosyl-*L*-methionine (SAM) as the methyl group donor and contain a conserved catalytic SET domain, apart from DOT1 (Martin and Zhang, 2005; Rea et al., 2000).

Prior to the discovery of KDMs, the histone methylation was widely considered to be irreversible, since the methyl group is quite stable, especially at the heterochromatin regions (Bannister et al., 2002). However, this idea was changed by the discovery of the first KDM, the amine oxidase-domain-containing mammalian protein LSD1 (lysine-specific demethylase 1), which could remove the methyl groups from the H3K4me2 and H3K4me3 (Shi et al., 2004). In addition, it has also been found that LSD1 could form a complex with androgen receptor (AR) and play a role in the removal of H3K9 methylation in mammals (Metzger et al., 2005).

Following LSD1, five other demethylases were identified and all these enzymes contain a common catalytic structure, the JmjC-domain. Among these enzymes, JHDM1 demethylates H3K36me1 and H3K36me2, whereas JHDM2A demethylates H3K9me1 and H3K9me2 (Klose et al., 2006; Tsukada and Zhang, 2006). However, the

tri-methylation state of these modified residues, including H3K36me3 and H3K9me3, could only be removed by JHDM3A or JMJD2A (Klose et al., 2006; Tsukada and Zhang, 2006). In addition, since no JmjC-domain has been found in the LSD1, the JmjC-domain demethylases must exert their function through a different mechanism from that of LSD1.

Unlike histone acetylation, histone methylation at lysine residues can play either a positive or a negative role in gene regulation. So far, there are six residues that have been well studied, including the lysine 4, 9, 27, 36 and 79 on the tail of histone H3 and lysine 20 on the tail of histone H4. In general, the methylation marks present at the H3K4, H3K36 and H3K79 sites are associated with gene expression, whereas the rest of them play a negative role in the gene expression. For instance, the H3K4 methylation, which was first identified in the vicinity of the chicken β-globin locus and at the active mating-type loci in budding yeast (Litt et al., 2001; Noma et al., 2001), is always associated with gene expression.

In mammals, H3K4me1 is present at enhancers, whereas H3K4me3 is present at the promoters of the transcriptionally active genes. Structurally, H3K4me can be recognized by either the tandem set of chromodomains or PHD finger containing proteins (Sims et al., 2006). However, other proteins without these domains, such as ISWI ATPase, may also bind to the H3K4 methylation in an indirect way. Mechanistically, there are three components that play a role in the establishment of the H3K4 methylation. First, the RNA Polymease II (Pol II) that containing serine-5 phosphorylation at its C-terminal domain could promote the establishment of H3K4 methylation by recruitment of SET1 to the promoters. In addition, via the interaction with the Pol II, the PAF complex could indirectly recruit the SET1 and coordinate the

setting of H3K4me3 (Zhu et al., 2005). Furthermore, the monoubiquitylation of H2B at lysine 123 in yeast (lysine 120 in human) also plays a role in the establishment of H3K4me3 (Dover et al., 2002; Sun and Allis, 2002).

H3K9 methylation is another well-characterized histone modification, since one of its modifiers, SUV39H1, was the first histone lysine methyltransferase to be identified (Rea et al., 2000). As the *Drosophila* homolog, Su(var)39, was initially identified as a suppressor of the position effect variegation, it was suggested that the SUV39 plays a negative role in gene transcription. In mammals, SUV39H1 and SUV39H2 methylate the lysine 9 on histone H3 to H3K9me3 state, which provides the binding sites for the chromodomain within the heterochromatin protein 1 (HP1) (Nielsen et al., 2001). However, HP1 not only binds to the H3K9me3, but also interacts with the SUV39, leading to the methylation of the lysine 9 within the neighbouring nucleosomes. Therefore, the heterochromatin patterns tend to be spread bi-directionally (Nakayama et al., 2001).

The presence of H3K9 methylation is often associated with DNA methylation. In the DNMTs-deficient cancer cells, apart from loss of DNA methylation globally, the level of the H3K9me3 is also reduced dramatically. This observation can be attributed to the fact that the stability of the histone methyltransferase SETDB1, which is required for the establishment of H3K9me3, depends on the presence of the methylated DNA binding protein 1 (MBD1) in mammalian cells (Martin and Zhang, 2005; Zhang and Reinberg, 2001). Unlike H3K9me3, which is predominantly present at the heterochromatin regions, H3K9me2 is present at the euchromatic regions and the nuclear periphery (Yokochi et al., 2009). In mammals, the main modifiers for depositing of the H3K9me2 are G9a and GLP. In addition, it has been found that the

G9a/GLP complex is the predominant form of these two enzymes that exist *in vivo* (Tachibana et al., 2005; Ueda et al., 2006).

H3K27 methylation is another repressive epigenetic mark, which is associated with gene silencing and X chromosome inactivation. The H3K27me3 modifier EZH2 is a component of the Polycomb repressive complex 2 (PRC2). Therefore, the H3K27me3 is associated with the Polycomb-mediated repression. In animals, the Polycomb group proteins are required for the repression of development-associated genes, including the genes within the *Hox* loci (Kirmizis et al., 2004). In addition, H3K27me3 is also associated with the malignant tumours, since the EZH2 is overexpressed in many types of the cancer cells (Martin and Zhang, 2005). Interestingly, histone demethylase JARID1D is thought to interact with the PRC2 and plays a role in the removal of the H3K4 methylation from the euchromatin regions (Pedersen and Helin, 2010). Apart from PRC2, Polycomb repressive complex 1 (PRC1) could also recognise H3K27me3 through its chromodomain and deposit ubiquitination at the H2AK119 site (Wang et al., 2004).

The H3K36 methylation is highly enriched on the gene body regions of the actively transcribed genes and promotes the efficient elongation. The RNA pol II that contains the serine-2 phosphorylation within its C-terminal domain could recruit the SET2, which is essential for the establishment of H3K36me3. In addition, H3K36me3 also plays a role in the suppression of the aberrant transcription in yeast (Carrozza et al., 2005; Joshi and Struhl, 2005). On one hand, H3K36me3 could recruit EAF3 and RPD3S (a homolog of mammalian HDAC2) to the gene body regions. Since RPD3S could later remove the histone acetylation, which is required for the initiation of transcription, the presence of the H3K36me3 at the gene body regions may prevent the

binding of Pol II to the cryptic promoters. On the other hand, since H3K36me3 also prevents the histone exchanges within the gene bodies (Venkatesh et al., 2012), lack of the active mark in the gene body regions may also prevent the initiation of the cryptic transcription.

Since the lysine 79 of histone H3 is located within the core of the nucleosome, H3K79me is an unusual mark. From yeast to eukaryotes, it has been found that the H3K79me is present on the coding regions of the active genes, which suggests that the H3K79me is a mark of active transcription (Fingerman et al., 2007). In budding yeast, H3K79me prevents the binding of repressive proteins to euchromatin, such as the SIR2 and SIR3 (Frederiks et al., 2008). H4K20me2 and H4K20me3 present at the pericentrometic heterochromatin regions and the levels of these marks are associated with the levels of H3K9me3. H4K20me1 plays a role in transcriptional repression and X inactivation (Karachentsev et al. 2005; Kohlmaier et al. 2004). Recent study has found H4K20me1 also enriched at the actively transcribed genes (Vakoc et al, 2006).

1.4.3 H3K9 methyltransferases

The catalytic domain of histone methyltransferases is SET domain, which stands for suppressor of variegation, enhancer of zeste and trithorax (Tschiersch et al, 1994). Actually, after SUV39H1 (KMT1A) was found as the first KMTs, proteins containing SET domain have been identified in various organisms, including viruses, bacteria, yeast, fungi and multi-cellular organisms and all these proteins have been shown to be able to methylate histone and non-histone substrates *in vivo* (Pontvianne et al., 2010; Reyes-Dominguez et al., 2008). In addition to SUV39H1 and SUV39H2, there are some other HKMTs that also play a role in the establishment of H3K9 methylation in mammalian cells, including G9a (EHMT2), GLP (EHMT1), SETDB1 (ESET), PRDM3

and PRDM16 (Pinheiro et al., 2012; Tachibana et al., 2001). Biochemically, both SUV39H1 and G9a can deposit mono-, di-, and trimethylation at the lysine 9 of histone H3 *in vitro* (Kubicek et al., 2007). However, investigations that carried out in the *Suv39h1* and *Suv39h2* double knockout or the *G9a*-deficient cells illustrate that the main function of SUV39H1 and SUV39H2 *in vivo* is to deposit H3K9me3 at the pericentromeric heterochromatin regions, whereas G9a plays a critical role in the establishment of H3K9me1 and H3K9me2 in euchromatin (Tachibana et al., 2002; Peters et al., 2003; Rice et al., 2003).

Similar to the SUV39H1 and SUV39H2, SETDB1 (ESET) is another HKMTs, which also plays a role in the establishment of H3K9me3 in mammalian cells. It has been found that SETDB1 is essential for the survival of the ES cells and plays an essential role in the development of the pre-implantation embryos (Dodge et al., 2004). In addition, in the *Setdb1* tissue specific knockout animals and ES cells, it was found that SETDB1 is also required for the maintenance of articular cartilage (Lawson et al., 2013a) and plays a role in the regulation of the osteoblast differentiation during the bone development (Lawson et al., 2013a; Lawson et al., 2013b).

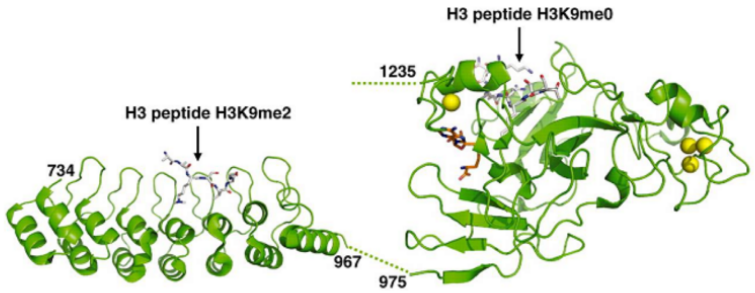
Originally described as the G9a-like protein, GLP has been found to possess the same substrates specificity to those of G9a *in vitro* (Tachibana et al., 2005). In mammals, although G9a and GLP share the same histone substrates and exert their methyltransferase activities independently, the levels of H3K9me1 and H3K9me2 are significantly reduced in either *G9a* or *Glp* knockout ES cells (Tachibana et al., 2005). Furthermore, since the levels of H3K9me1 and H3K9me2 in the *G9a* or *GLP* single knockout or the double knockout cells are comparable, it has been suggested that G9a can not compensate for the functions of GLP and *vice versa* (Tachibana et al., 2005).

Furthermore, it has been found that G9a and GLP can form either homodimers or heterodimers *in vivo* via the interactions of their SET domains. In mouse ES cells, the endogenous G9a and GLP exist predominantly as a stoichiometric G9a/GLP heterodimer and a C2H2 zinc finger protein WIZ serves as another stable component in the G9a/GLP complex (Tachibana et al., 2005; Tachibana et al., 2008) (Figure 1.7). In addition, the NHLC motif located in the pre-SET domain of either G9a or GLP is essential for the dimerization of these two enzymes (Figure 1.7) and the mutations within this motif cause both reduction of H3K9me2 and loss of DNA methylation at the G9a targeted loci (Tachibana et al., 2008).

A

Name	Transcript ID	Length	Protein
Ehmt2-001	OTTMUST00000039149	4026 bp	1263 aa
Ehmt2-002	OTTMUST00000039148	3934 bp	1229 aa
Ehmt2-003	OTTMUST00000039150	3882 bp	1206 aa
Ehmt2-004	OTTMUST00000039151	3746 bp	1172 aa
Ehmt2-009	OTTMUST00000039156	1325 bp	442 aa
Ehmt2-014	OTTMUST00000096155	583 bp	194 aa

B



C

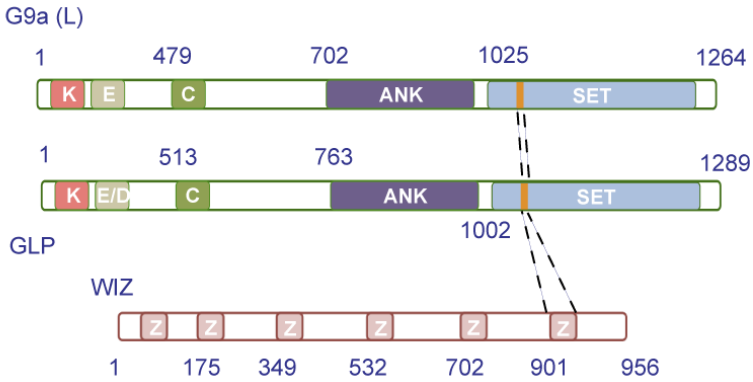


Figure 1.7 The transcripts and structure of mouse G9a and the core components of G9a/GLP complex *in vivo*.

(A) In mouse, there are six different protein coding transcripts of G9a. Data from http://vega.sanger.ac.uk. (B) The structure of ANK and SET domains within G9a protein (adapted from Collins and Cheng, 2010). G9a recognises and binds to H3K9me1/2 via the fourth and fifth alpha-loop within its ANK domain, whereas through its SET domain, G9a is able to bind to the unmethylated histone H3 peptide and establish H3K9me1/2. (C) G9a/GLP/Wiz tri-molecular complex is the predominant form of G9a exists *in vivo*. G9a binds to G9a-like protein (GLP) via its SET domain and the sixth zinc finger within Wiz is involved in the interaction with G9a/GLP complex. In structure, both G9a and GLP contain the ANK and SET domains. The orange bars within G9a and GLP represent for the amino acid motifs (NHLC 1165-1168 in G9a and NHHC 1198-1201 in GLP) that are required for the dimerization of G9a and GLP. ANK: ankryin repeat domain; SET: a catalytic methyltransferase domain; Z: C2H2 zinc finger.

Moreover, it has been found that the ankyrin repeat (ANK) domains within G9a and GLP are able to recognise and bind to the H3K9me1 and H3K9me2 through the amino acids within the fourth and the fifth ankyrin repeat loop, including W839, W844, W877 and E847 (Collins et al., 2007). In addition, among these amino acids, any mutation could completely disrupt the interactions between the G9a/GLP complex and H3K9 methylation (Collins et al., 2007).

Apart from G9a and GLP, PRDM3 and PRDM16 also play a role in the establishment of H3K9me1 in mouse embryonic fibroblasts (MEFs) (Pinheiro et al., 2012). Interestingly, apart from exists within the nucleosomes, the non-nucleosomal form of H3K9me1 also presents in the cytoplasm and nucleus. This makes H3K9me1 different from the other histone methylations, since all the other histone methylations cannot be detected in a non-nucleosomal manner (Loyola et al., 2006). However, apart from H3K9me1, PRDM3 and PRDM16 also contribute to the H3K9me3, since in the *Prdm3* and *Prdm16* knockout cells, H3K9me3 cannot be fully established at the major satellite regions (Pinheiro et al., 2012).

1.4.4 H3K9 methylation and DNA methylation

The first evidence that indicating H3K9 methylation is required for DNA methylation was reported in the fungus *Neurospora crassa* (Tamaru and Selker, 2001). In *Neurospora*, DIM-5 functions as the main KMT, which is essential for the establishment of H3K9me3. Since the heterochromatin protein 1 (HP1) is able to recognise and bind to the H3K9me3, which can further recruit DNA methyltransferase (DIM-2) to the H3K9me3-marked regions, the presence of H3K9me3 is required for the *de novo* DNA methylation in *Neurospora* (Tamaru and Selker, 2001; Rountree and Selker, 2010).

In plants, H3K9 methylation is also associated with DNA methylation (Johnson et al., 2007). In *Arabidopsis thaliana*, the histone methyltransferase KRYPTONITE (KYP) is able to directly bind to the methylated DNA and establish H3K9me2 at the targeted loci. Since the newly established H3K9me2 can further recruit the *de novo* DNA methyltransferase CMT3 to the chromatin, DNA methylation and H3K9 methylation mutually reinforce each other.

The methylation of H3K9 is also linked to DNA methylation in mammals. It has been found that DNA methylation that presents at the major, but not minor, satellite repeat region is reduced in the *Suv39h1* and *Suv39h2* double knockout ES cells (Lehnertz et al., 2003). Moreover, since H3K9me3 is required for the recruitment of HP1α and HP1β and these HP1 proteins are able to further recruit DNMT3B to the chromatin, the DNA methylation that presents at the pericentromeric heterochromatin region is partially dependent on the presence of SUV39H1/2-mediated H3K9me3.

Additionally, it has also been found that both H3K9me2 and the global DNA methylation are reduced in the *G9a* or *Glp* knockout ES cells (Dong et al., 2008; Epsztejn-Litman et al., 2008; Tachibana et al., 2008). Interestingly, since the loss of DNA methylation in the *G9a-/-* ES cells can be restored partially by the catalytically inactive G9a, it has been found that G9a promotes DNA methylation independently of its catalytic activity (Dong et al., 2008). Actually, G9a is able to directly interact with the DNMTs in mammalian cells. For instance, the N-terminal domain of GFP-tagged G9a directly interacts with the his-tagged DNMT1 in Cos-7 cells (Esteve et al., 2006) (Figure 1.8), whereas in human colon cancer cells, G9a forms a complex with DNMT1 (Fraga et al., 2004). Furthermore, in *Hela* cells, the chromatin binding patterns of G9a is correlated with the chromatin positioning of DNMT1 (Esteve et al., 2006).

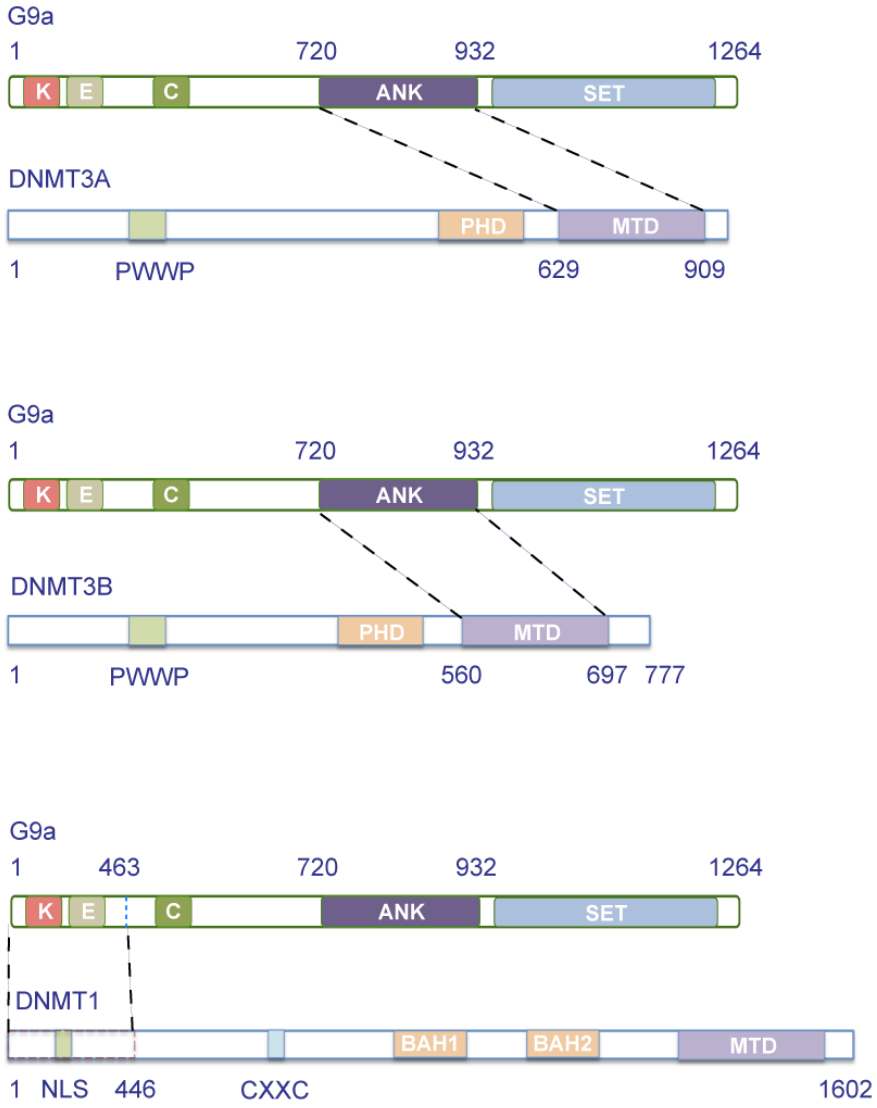


Figure 1.8 The schematic view of interactions between G9a and DNMTs.

In vivo, G9a can interact with the DNA methyltransferases, which play a role in either the maintenance or *de novo* DNA methylation. The ANK domain of G9a and the MTD domains of DNMT3A/B are essential for the interaction between G9a and DNMT3s. Both the N-terminus of G9a and DNMT1 are involved in the interaction of these two enzymes. ANK: ankyrin repeats domain; SET: SET domain; MTD: methyltransferase domain; PHD: Plant Homeo Domain; PWWP: Pro-Trp-Trp-Pro motif; NLS: nuclear localization signal; BAH1 and 2: bromo-adjacent homology 1 and 2; CXXC: a zinc finger domain which binds to unmethylated DNA substrate. K: lysince rich region; E: Glutamic acid rich region; C: Cysteine rich region; E/D: Glutamic acid and Aspartic acid rich region. Figures are adapted from published articles Esteve et al., 2006 and Epsztejn-Litman et al., 2008.

By contrast, in mouse ES cells, the chromatin binding patterns of G9a is independent of the positioning of DNMTs on the chromatin (Oda et al., 2006). Moreover, it has been found that both DNMT3A and DNMT3B are able to directly interact with G9a/GLP complex *in vivo* (Epsztejn-Litman et al., 2008). By co-expression of the truncated forms of G9a and DNMT3A in 293 cells, it was reported that the ANK domain within G9a is essential for the interaction between G9a and the methyltransferase domain (MTD) of DMNT3A (Epsztejn-Litman et al, 2008) (Figure 1.8). In addition, the hemi-methylated DNA binding protein UHRF1 was also found to interact with DNMT1 and H3K9 methylation. During DNA replication, UHRF1 forms a complex with DNMT1 (Sharif et al., 2007) and binds to the H3K9me2 and H3K9me3 via its tandem tudor domain and the plant homeo-domain (PHD) (Rottach et al., 2010).

1.4.5 H3K9 methylation and imprinted DNA methylation

Previous studies have found that H3K9me3 is enriched not only at the ICRs on the methylated allele, but also at the promoters of several imprinted genes, including *H19* and *Cdkn1c* (Ager et al., 2008). Recently, Quenneville *et al* reported that the zinc finger protein ZFP57 and its cofactor KAP1 are able to selectively bind to the H3K9me3-marked methylated allele at the ICRs and play an important role in the maintenance of the imprinted DNA methylation in ES cells (Quenneville et al., 2011).

G9a is also essential for the maintenance of DNA methylation at the imprinted loci in mammals. In mouse *G9a-/-* ES cells, H3K9me2 is reduced at the PWS associated ICR. In addition, one of the imprinted genes that located within this imprinting cluster, named *Snrpn*, is bi-allelically expressed in the *G9a-/-* ES cells (Xin et al., 2003). However, it has been reported that in the *Dnmt1-/-* ES cells, H3K9me2 alone is sufficient for the maintenance of the monoallelic expression of *Snrpn* (Esteve et al.,

2006). Taken together, these observations indicate that the absence of imprinted DNA methylation is not sufficient to cause biallelic expression of *Snrpn* in mouse ES cells. Furthermore, G9a has also been found to be responsible for the allele-specific repression of imprinted genes in the placenta, particularly the genes that are imprinted in the trophoblast, but not in the embryos (Nagano et al, 2008; Wagschal et al., 2008). However, since the imprinted genes that lost imprinting in the absence of G9a were not dependent on the DNA methylation, the role for G9a in the maintenance of genomic imprinting in the placenta seems requires the heterochromatin marks, such as H3K9me2 or H3K9me3.

1.5 Patterns of promoter DNA methylation in *G9a*-/- ES cells

It has been shown that in the *G9a-/-* ES cells, some genomic regions display hypomethylation, including the major satellite repeats, endogenous retroviruses and promoters of a few protein coding genes (Shinkai and Tachibana, 2011)However, the extent to which the DNA methylation is lost in the absence of *G9a* was still unknown. Previous work in the lab has shown that approximately 50% of the global DNA methylation is lost in the *G9a-/-* ES cells when compared to the wild-type ES cells (Myant et al., 2011). In addition, the promoters of a number of the protein coding genes, including the trophectoderm-specific *Elf5* and *Rhox* genes, display detectable hypomethylation (Myant et al., 2011).

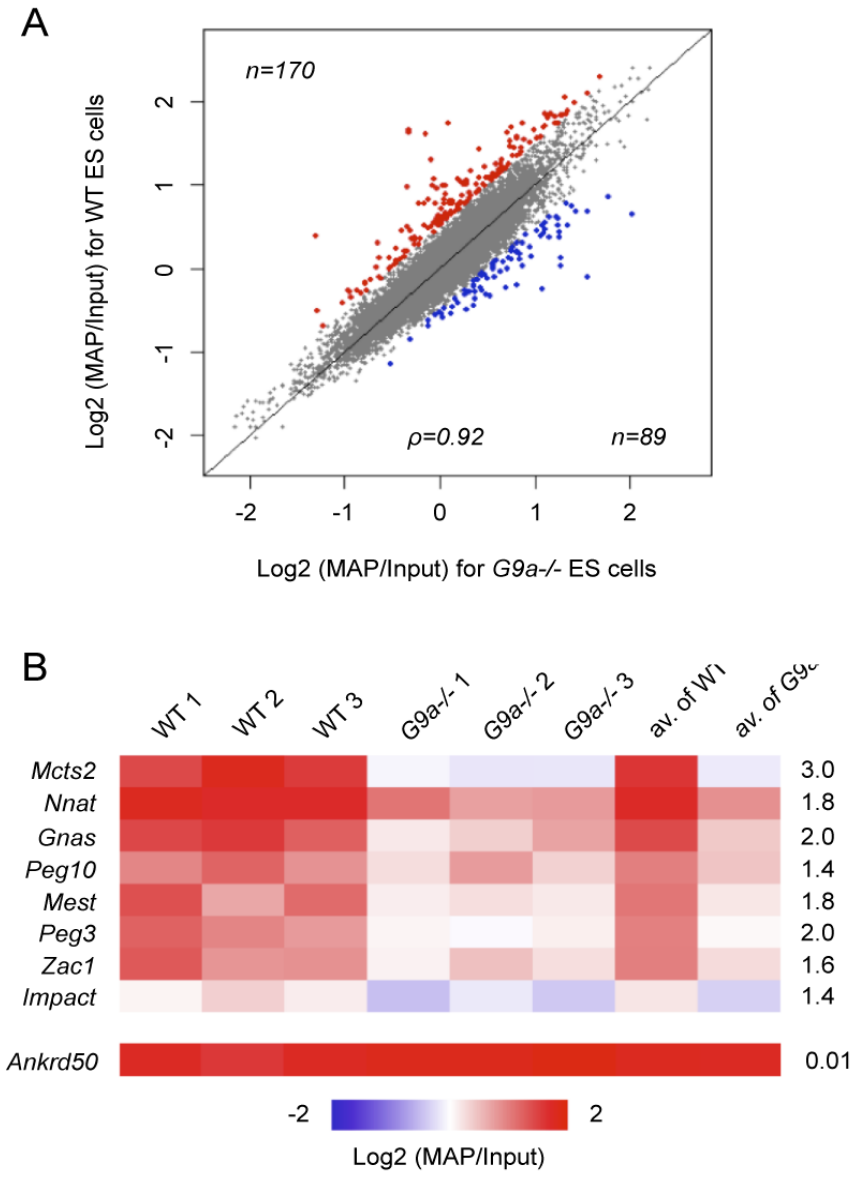


Figure 1.9 Loss and gain DNA methylation at promoters in *G9a-/-* ES cells.

(A) Scatter plots comparing average DNA methylation values of Log2 (MAP/Input) from replicate microarrays for all 1kb promoter regions in wild-type (y-axis) versus *G9a-/-* ES cells (x-axis). The red and blue spots indicate the promoters showing more than 1.4-fold changes of loss or gain of DNA methylation in *G9a-/-* ES cells. "n" indicates the number of promoters displaying loss or gain of DNA methylation. "ρ" is the Spearman correlation coefficient. (B) Fold changes of DNA methylation at promoters of eight imprinted genes in wild-type and *G9a-/-* ES cells. *Ankrd50* is a positive control gene. av. of WT and av. of *G9a-/-* represent hte average fold changes of three times of Log2 (MAP/Input) values in wild-type and *G9a-/-* ES cells, respectively. Data from Ausma Termanis, unpublished.

To investigate whether *G9a* is required for DNA methylation at the promoters of protein coding genes, Ausma Termanis, a former PhD student in the lab, carried out affinity purification of methylated DNA from wild-type and *G9a-/-* ES cells. The purified methylated DNA was then hybridized to the microarrays, which contains the high density of the oligonucleotide within the promoters of all the annotated mouse protein coding genes. These experiments identified 170 hypo- and 89 hyper-methylated promoters in the *G9a-/-* ES cells when compared with the wild-type control (Figure 1.9A) (Ausma Termanis, unpublished). Strikingly, the promoters of some maternal imprinted genes were identified among the hypo-methylated loci (Figure 1.9B) (Ausma Termanis, unpublished). This interesting observation suggests that the histone methyltransferase G9a may play a role in the protection of the imprinted DNA methylation in mouse ES cells.

1.6 Aims of the project

It has been shown that loss of the DNA methylation at ICRs leads to the aberrant expression of the imprinted genes. Therefore, it has been found that the imprinted genes are directly regulated by the imprinted DNA methylation. In addition, since once established, the patterns of the imprinted DNA methylation are maintained stably throughout life and numerous human disorders have been found to be associated with the abnormal expression of the imprinted genes, it is important to investigate how the imprinted DNA methylation is maintained in mammals.

Previous observations have suggested that apart from depositing H3K9me2, the histone methyltransferase G9a and its partner GLP also affect the global levels of DNA

methylation in mammalian cells (Dong et al., 2008). In addition, G9a is also involved in the maintenance of imprinted DNA methylation at some imprinted loci (Xin et al., 2003). Furthermore, the chromatin binding protein STELLA requires the G9a associated H3K9me2 in the maintenance of DNA methylation during the early developmental stage (Nakamura et al., 2007; Nakamura et al., 2012). Together with the initial observation suggesting that maternally derived ICRs are hypomethylated in *G9a-/-* ES cells (Ausma Termanis, unpublished), these evidence led to the hypothesis that G9a may play an important role in the protection of the imprinted (and also the non-imprinted) DNA methylation in mouse embryonic stem (ES) cells from demethylase activities.

Therefore, the first aim of my project was to validate independently the microarray data and investigate whether G9a is required for the DNA methylation at some or all the ICRs in ES cells. As paternal ICRs, including *Igf2/H19, Dlk1/Dio3* and *Rasgrf1*, are located at the intergenic regions, while maternal ICRs are at the promoters, I also would like to determine whether both maternal and paternal ICRs lose DNA methylation in the *G9a-/-* ES cells. Since the imprinted DNA methylation is established in the primordial germ cells (PGCs) and is normally stably maintained throughout life (Lucifero et al., 2002), the third aim of my project was to investigate whether and how does G9a maintain the imprinted DNA methylation in ES cells. Given that G9a serves as the main modifier for depositing H3K9me1 and H3K9me2 in ES cells and the previous studies showed that G9a protein, but not its catalytic activity, affects the patterns of DNA methylation in ES cells (Dong et al., 2008), I also aimed to investigate whether the catalytic activity of G9a is essential for the protection of the imprinted

DNA methylation and, if so, to determine how does G9a protect the DNA methylation at the ICRs in ES cells.

Chapter 2 Materials and Methods

2.1 Materials

2.1.1 Buffers and solutions:

Phosphate buffered saline (PBS): 140 mM NaCl, 3 mM KCl, 2 mM KH_2PO_4, 10 mM Na_2HPO_4. Stored at room temperature.

Tris-EDTA (TE): 10 mM Tris pH 7.5, 1 mM EDTA pH 8.0.

Tris-acetate EDTA (TAE): 40 mM Tris, 20 mM glacial acetic acid, 1 mM EDTA and adjusted to pH 8.0.

Bisulfite conversion solution: 3.8 g sodium bisulfite was dissolved in 5 ml H_2O and 1.5 ml freshly made 3 M NaOH in the dark. 110 mg hydroquinone was dissolved in 1 ml water at 55 °C for 10 minutes and subsequently added to the sodium bisulfite solution.

Special PCR buffer: 166 mM $(NH4)_2SO_4$, 670 mM Tris pH 8.0 and 100 mM Beta-mercaptoethanol.

2.5 × Sequencing buffer: 20 mM Tris-HCl pH 8.0, 5 mM $MgCl_2$.

Protein manipulation solutions:

Hypotopic NE1 buffer: 20 mM HEPES pH7.0, 10 mM KCl, 1 mM $MgCl_2$, 0.1% (v/v) Triton X-100, 20% (v/v) glycerol, 0.5 mM DTT and complete protease inhibitors. DTT and complete protease inhibitors were added immediately prior to use. Stored at 4°C.

SDS-PAGE loading buffer (5 ×): 0.5 M Tris pH 6.8, 10% (v/v) glycerol, 2% SDS, 0.05% bromophenol blue, 250 mM DTT. Stored at room temperature, DTT added just prior to use.

SDS-PAGE separating gel: 0.1% SDS, 0.05% ammonium persulfate, desired concentration of acrylamide and 375 mM Tris pH 8.8 and made up to 10 ml with distilled H_2O.

SDS-PAGE stacking gel: 0.1% SDS, 0.05% ammonium persulfate, 4% acrylamide, 125 mM Tris pH 6.8 and made up to 10 ml with distilled H_2O.

Western blotting running buffer: 25 mM Tris, 250 mM Glycine, 0.1% SDS.

Western blotting transfer buffer: 25 mM Tris, 250 mM Glycine.

Ponceau S staining solution: 1% (v/v) glacial acetic acid, 0.5% (w/v) Ponceau S.

TBS solution: 50 mM Tris-Cl, pH 7.6; 150 mM NaCl

MeDIP solutions:

10 × IP buffer: 100 mM NaH_2PO_4 pH 7.0, 1.4 M NaCl, 0.5% Triton X-100

Digestion buffer: 50 mM Tris pH 8.0, 10 mM EDTA, 0.5% SDS.

ChIP solutions:

10 × cross-linking solution: 500 mM HEPES pH 7.9, 1.5 M NaCl, 10 mM EDTA, 5 mM EGTA.

L1 buffer: 50 mM HEPES pH 7.9, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100.

L2 buffer: 10 mM Tris pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA.

L3 buffer: 10 mM Tris pH 8.0, 1 mM EDTA, 0.5 mM EGTA.

ChIP dilution buffer: 20 mM Tris pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100.

Wash buffer 1: 20 mM Tris pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS.

Wash buffer 2: 20 mM Tris pH 8.0, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS.

Equipment: Quantitative PCR instrument (Lightcycle 480, Roche), Odyssey scanner (Licor), Biorupter (Diagenode), PCR instruments (Biometra), CO_2 incubator (Thermo Scientific).

2.1.2 Antibodies

anti-G9a (Mouse monoclonal Perseus Proteomics PP-A8620A-00): 1:1000 working concentration in 4% milk + 0.1% Tween 20.

anti-GLP (Mouse monoclonal Perseus Proteomics PP-B0422-00): 1:1000 working concentration in 4% milk + 0.1% Tween 20.

anti-HDAC1 (Rabbit polyclonal Santa Cruz sc-7872/D171): 1:1000 working concentration in 4% milk + 0.1% Tween 20.

anti-DNMT3a (Mouse monoclonal Abcam ab13888): 1:1000 working concentration in 4% milk + 0.1% Tween 20.

anti-DNMT3b (Mouse monoclonal Abcam ab13604): 1:1000 working concentration in 4% milk + 0.1% Tween 20.

anti-DNMT1 (Rabbit polyclonal Santa Cruz sc-20701/H300): 1:1000 working concentration in 4% milk + 0.1% Tween 20.

anti-H3K9me1 (Rabbit polyclonal Upstate 07-450): 1:2000 working concentration in 4% milk + 0.1% Tween 20.

anti-H3K9me2 (Mouse monoclonal Abcam ab1220): 1:2000 working concentration in 4% milk + 0.1% Tween 20 and 1:2000 working concentration for ChIP.

anti-H3K9me3 (Rabbit polyclonal Upstate 07-442): 1:2000 working concentration in 4% milk + 0.1% Tween 20.

anti-H4 (Rabbit polyclonal Upstate 06-866): 1:500 working concentration in 4% milk + 0.1% Tween 20.

anti-Ran (Mouse monoclonal Abcam ab11693): 1:5000 working concentration in 4% milk + 0.1% Tween 20.

anti-Stella (Rabbit polyclonal Santa Cruz sc-67249/M150): 1:1000 working concentration in 4% milk + 0.1% Tween 20.

anti-LSH (Mouse monoclonal Santa Cruz sc-46665/H0205): 1:500 working concentration in 4% milk + 0.1% Tween 20.

anti-5meC (Mouse monoclonal Eurogentec BI-MECY-1000): 1:2000 working concentration for MeDIP.

IRDye 800CW Donkey anti-Mouse IgG: 1:10000 working concentration in 4% milk + 0.1% Tween 20.

IRDye 800CW Donkey anti-Rabbit IgG: 1:10000 working concentration in 4% milk + 0.1% Tween 20.

IRDye 680RD Donkey anti-Rabbit IgG: 1:10000 working concentration in 4% milk + 0.1% Tween 20.

2.1.3 Primers

All primers were synthesised from MWG Eurofins. Lyophilised primers were

diluted in distilled H_2O to 100 μM stock concentration and stored at -20 ^oC. The final concentration of the primers for the normal PCR and qPCR is 5 μM and 2.5 μM, respectively.

Primers for bisulfite DNA sequencing:

Ankrd50	mAnkrd50-BS-F1 (out)	GGATGTGGTGGATTTGTTGTTA
	mAnkrd50-BS-F2 (in)	TTGTTGTTAGAAGGAGGAGTAGATGT
	mAnkrd50-BS-R	TCCAAACCTCTATCCAAAAAAATAC
Igf2r	mIgf2r-BS-F (out)	TTAGTGGGGTATTTTTATTTGTATGG
	mIgf2r-BS-F (in)	GTGTGGTATTTTTATGTATAGTTAGG
	mIgf2r-BS-R	AAATATCCTAAAAATACAAACTACAC
Snrpn	mSnrpn-BS-F (out)	GGGTTGTTAAAAATTTTAATAAGTTTAAA
	mSnrpn-BS-F (in)	TTTAGAATGTTTTGGTTAAATAGGATGTA
	mSnrpn-BS-R	AAAAAAACAAAAACCCCTACATTAC
H19	mH19-BS-F (out)	GTTTGAGGAGTTTTAAGGTAGAAGG
	mH19-BS-F (in)	TTTTGTTGAATTTGGGGTATTTAAA
	mH19-BS-R	TCACTCAAACATAACATTCAATAATTC
Zac1	mZac1-BS-F	ATTTGTTATTTAGTTTGGGTTGGG

	mZac1-BS-R1 (in)	CCCAAATTCAAAATTTATCACCTC
	mZac1-BS-R2 (out)	ATTCTCCCAAAAATTCTTAAAAATC

Primers for MeDIP or ChIP:

Dpep3	IP-Dpep3-F	GAAGTAACACCCCCAGCAGGGACA
	IP-Dpep3-R	CCCTGAATCGAAGGTCCGAACCCA
Ankrd50	IP-Ankrd50-F	TCTCGTCCAAGCCTCTGTC
	IP-Ankrd50-R	GTCGATCACACCGATAACAAC
Igf2r	IP-Igf2r-F	CGTGATCCTTGGTTGTGCTGAG
	IP-Igf2r-R	CCAACCGGAATCGCATTAAAACC
Snrpn	IP-Snrpn-F	CAGGACATTCCGGTCAGAG
	IP-Snrpn-R	TACTAGAATCCACAAGCCCAG
H19	IP-H19-F	GGTGGCAGCATACTCCTATAT
	IP-H19-R	CTCGGCAACTTCGGTCTTAC
Zac1	IP-Zac1-F	GCATCTGCGATTTGTCACTC
	IP-Zac1-R	CTTGCTCTCCAGTCCCGATA
Peg3	IP-Peg3-F	CAGAGGACCCTGACAAGGAG

	IP-Peg3-R	AGCACAGCACTCTACGCACA
Peg10	IP- Peg10-F	TCCTGACCAACTACGACCTG
	IP- Peg10-R	CCATACTCACCACACGAGGA
Peg13	IP- Peg13-F	AGCTGAGCGAACCCCTTTAC
	IP- Peg13-R	CGCAGGTCTTCTATCCAACC
Kcnq1ot1	IP-Kcnq1ot1-F	CAGCACGGATCACTCCAG
	IP-Kcnq1ot1-R	AAAGCTCTCCAAGTAGAATCACA
Rasgrf1	IP-Rasgrf1-F	CTGCACTTCGCTACCGTTTC
	IP-Rasgrf1-R	AGTAGCAGTCGTGGTAGTTG
GnasXL	IP-GnasXL-F	CACTGAGACCTGCGTCCTCT
	IP-GnasXL-R	TGGTCGGCCAACAACTTTAG
Ndufa	IP-Ndufa-F	TCCGCACCGTTACTCGCACG
	IP-Ndufa-R	AGCCACCGTCGCTTCCTCCT

Primers for molecular cloning:

G9a	G9a-BglII-F	CGGTAGAGATCTGTCATGTGCTGCTGCA GGC
	G9a-BglII-R	GGGAACCGTAGATCTTGGTGCAGCATG

		AAAAC
G9a	G9a-shMUT-F	GATTTAGATAACAAGGATGGTGAGGTC TACTGTATTGATGCCCGTTACTATGGC
	G9a-shMUT-R	GCCATAGTAACGGGCATCGATGCAATA AACTTCGCCATCCTTGTTATCTAAATC
G9a	G9a-shMUT2-F	CGATTTAGATAACAAGGATGGTGAAGT CTATTGTATCGACGCCCGTTACTATGGC
	G9a-shMUT2-R	GCCATAGTAACGGGCGTCGATACAATA GACTTCACCATCCTTGTTATCTAAATCG
G9a	G9a-AflII-F	GCCATCCTTAAGCGGGAGACCATGCGG CC
	G9a-BspEI-R	TCCCCAGTCCGGATGTCCCTGGAGC
G9a	G9a-ANKdel-F	GCCTTGGTCATCCTGACCCCAGAGCGCT CTGATGTGTGGTTTGC
	G9a-ANKdel-R	GCGCTCTGGGGTCAGGATGACCAAGGC TTTTTCCAGGGCTTCCC
G9a	G9a-WA-EA-F	GAATCACATCGATGTGCTTGTGAGCGG CTGCAGCGATGATGGGCGTCCAGCCCC CACTGTCCTG

	G9a-WA-EA-R	CAGGACAGTGGGGGCTGGACGCCCATC ATCGCTGCAGCCGCTCACAAGCACATC GATGTGATTC
G9a	G9a-AflII-F	GCCATCCTTAAGCGGGAGACCATGCGG CC
	G9a-BspEI-R-out	CCTCCGCTGAATGCTTGCACTTCTCAGA GCC
G9a	G9a-NH-LE-F	CTATGGCAACATCAGCCGATTCATTCTC GAGCTGTGTGACCCCAACATCATCCC
	G9a-NH-LE-R	GACAGGGATGATGTTGGGGTCACACAG CTCGAGAATGAATCGGCTGATGTTGCC
G9a	G9a-Seq-F	AGGTACGGCTGCAGCTCTACCG

Primers for quantitative reverse transcription PCR (qRT-PCR):

G9a	WT-G9a-F1	AGTGTAACCAGGCATGCTCC
	WT-G9a-R1-5-mut	TGCAGTAAACCTCGCCATCC
G9a	WT-G9a-F2-5-mut	AGGATGGCGAGGTTTACTGC
	WT-G9a-R2	GTCACCGTAGTCAAAGCCCA
G9a	WT-G9a-F3-6-mut	GGTTTACTGCATTGATGCCCG

	WT-G9a-R3	TCGGTCACCGTAGTCAAAGC
G9a	WT-G9a-F4-6-mut	TGGCGAGGTTTACTGCATTG
	WT-G9a-R4	CCCAGAATCGGTCACCGTAG
G9a	mG9a-QRT-F	ACACGGCATGGGATCTGACCCC
	mG9a-QRT-R	ACGGCTCCCCATCCACACCATT

Primers for DNA sequencing:

pJET	Seq-pJET-F	GCCTGAACACCATATCCATCC
G9a	Seq-G9a-MUT-F	GGACGGGCGGCTGCTCCAGGAGTTTAAC
G9a	Seq-G9a-MUT-R	CCGGGTGGGGGTCCAGCCGGGCCAGGCG
G9a	Seq-G9a-F	AGGTACGGCTGCAGCTCTACCG
G9a	Seq-G9a-R	CTGGGAGGGGAGACAAGGAAC

2.1.2 Embryonic stem (ES) cell lines

G9a knockout ES cell line: Generated by gene conversion of heterozygous *G9a+/-* ES cells through stringent antibiotic selection (Yoichi Shinkai).

G9a knockout wild-type G9a rescued ES cell line: G9a knockout ES cells were selected by 1 μg/ml puromycin after transfection with a linearized plasmid expressing Flag-tagged wild-type G9a (Yoichi Shinkai).

G9a knockout catalytically inactive G9a rescued ES cell line: G9a knockout ES cells were selected by 1 μg/ml puromycin after transfection with a linearized plasmid expressing catalytically inactive (H1092A) G9a (Yoichi Shinkai).

G9a knockdown ES cell lines: Short hairpin RNA knockdown of G9a in TT2 or E14 ES cells through 1 μg/ml puromycin selection (Tuo Zhang).

Exogenous wild-type G9a (short hairpin RNA resistant) expression and endogenous G9a knockdown ES cell line: Wild-type ES cells were selected by 1 μg/ml of puromycin and 200 μg/ml of hygromycin after transfection with a linearized plasmid expressing a wild-type short hairpin RNA resistant form of G9a and a linearized plasmid expressing short hairpin RNA (Tuo Zhang).

Exogenous ANK domain mutant (ANKm) G9a (short hairpin RNA resistant) expression and endogenous G9a knockdown ES cell line: Wild-type ES cells were selected by 1 μg/ml puromycin and 200 μg/ml of hygromycin after transfection with a linearized plasmid expressing an ANK domain mutant (W844A, E847A) short hairpin RNA resistant form of G9a and a linearized plasmid expressing short hairpin RNA (Tuo Zhang).

Exogenous ANK domain deleted (ANKdel) G9a (short hairpin RNA resistant) expression and endogenous G9a knockdown ES cell line: Wild-type ES cells were selected by 1 μg/ml puromycin and 200 μg/ml of hygromycin after transfection with by a linearized plasmid expressing an ANK domain deleted (Δ692-943aa) short hairpin RNA resistant form of G9a and a linearized plasmid expressing short hairpin RNA (Tuo Zhang).

Exogenous dimerization mutant ANKm (dmANKm) G9a (short hairpin RNA resistant) expressing and endogenous G9a knockdown ES cell line: Wild-type ES cells were selected by 1μg/ml of puromycin and 200 μg/ml hygromycin after transfection with a linearized plasmid expressing an ANK domain mutant (W844A, E847A) and dimerization mutant (N1165L, H1166E) short hairpin RNA resistant form of G9a and a linearized plasmid expressing short hairpin RNA (Tuo Zhang).

Exogenous dimerization mutant ANKdel (dmANKdel) G9a (short hairpin RNA resistant) expressing and endogenous G9a knockdown ES cell line: Wild-type ES cells were selected by 1μg/ml puromycin and 200 μg/ml of hygromycin after transfection with a linearized plasmid expressing an ANK domain deleted (Δ692-943aa) and dimerization mutant (N1165L, H1166E) short hairpin RNA resistant form of G9a and a linearized plasmid expressing short hairpin RNA (Tuo Zhang).

2.2 Methods

2.2.1 Molecular cloning

Polymerase chain reaction (PCR): polymerase chain reaction was performed in 1 × buffer IV containing 12 mM $MgCl_2$, 2.5 mM dNTPs, 7% (v/v) DMSO, polymerase (2.5 units in 50 μl), 5 μM primers and desired amount of template. Denaturation, annealing and elongation were dependent on application. For bisulfite sequencing, the buffer IV was replaced by the special PCR buffer. The results of PCR were assessed by agarose gel electrophoresis.

Restriction enzyme digestion: Restriction enzyme digests were carried out in a suitable volume according to the manufacturer's instructions. Incubation times varied from 2 hours to overnight at desired temperature, typically 37 ^{o}C. Completion of

digestion was assessed by agarose gel electrophoresis.

Gel extraction: the gel slice containing the desired DNA fragments was dissected and melted at 55 °C for 10 minutes. Then the DNA was extracted following the manufacturer's instruction using the GeneJET Gel Extraction Kit (K0692, Thermo Scientific).

Ligation (for general cloning): Purified DNA fragments were blunted in the 1 × ligation buffer containing system at 70 ^{o}C for 5 minutes and the pJET1.4 linearized plasmid backbone and T4 DNA ligase were added. The ligation reaction was performed at 16 °C for 12-16 hours and the ligation results were assessed by competent cell transformation and selection.

Plasmid Precipitation: Plasmids were purified from 5 ml or 500 ml overnight (12-14 hours) E.coli cultures according to manufacturer's instruction by using the Quiagen Miniprep or Maxiprep Kits, respectively. Plasmid DNA was eluted in desired volume of 1 × TE buffer and kept at -20 °C.

2.2.2 Sodium bisulfite DNA sequencing

Genomic DNA extraction: Cells cultured in T175 flask were resuspended in 1 ml of 1 × TE buffer. Proteinase K and SDS were added to a final concentration of 200 μg/ml and 1% respectively. Lysed cells were digested by Proteinase K at 55 °C overnight. RNA was degraded by 10 μl of 100 μg/ml RNase incubated at 37 °C for 4 hours. The salt concentration was subsequently adjusted to 200 mM by adding 40 μl of 5 N NaCl. Digested peptides were removed through two rounds of phenol:chloroform:isoamyl alcohol extractions followed by chloroform extraction. Genomic DNA was precipitated in 1 volume isopropanol containing 1/10 volume of 3 M NaOAc (pH 5.4) solution.

Finally, the DNA pellet was washed by 70% (v/v) ethanol and dissolved in a desired volume of 1 × TE buffer.

Bisulfite conversion of genomic DNA: 2 μg of genomic DNA was resuspended in 1 × TE buffer with a total volume of 25 μl. DNA was denatured at 110 °C for 5 minutes and chilled immediately on ice. 2.5 μl of freshly made 3 M NaOH was added and then the samples were incubated at 37 °C for 20 minutes. Finally, 270 μl of sodium bisulfite solution (see solutions section) was added to each sample. Overlaid with mineral oil, the samples were incubated overnight for sulfonation at 55 °C.

Isopropanol precipitation and desulfonation: The treated genomic DNA was precipitated in isopropanol and sodium acetate solution with 2.5 μg of glycogen. DNA pellets were resuspended in 25 μl of 1 × TE buffer and desulfonated by 3 M NaOH for 15 minutes at 37 °C. Samples were then precipitated by 32.5 μl of 5 M ammonium acetate (pH 7.0) and 180 μl 100% ethanol. Finally, DNA pellet was dissolved in 25 μl of 1 × TE buffer.

PCR amplification of interest regions: Design nested PCR primers using Methprimer software aiming for amplification regions across CpG di-nucleotides rich region and PCR products around 300-400 base pairs long. Before setting up the PCR reactions, denature the DNA at 100 °C for 10 minutes and place it on ice. Use 1-2 μl of sodium bisulfite treated DNA as template for the first round PCR reaction (50 μl system). For the nested PCR (50 μl system), take 2 μl of the first round PCR product as template for the PCR reaction. The following program was used for both the first round and the nested PCR reactions: 95 °C for 5 minutes and 22 (first round) or 30 (nested) cycles of cycling with each cycle at 95 °C for 30 seconds, appropriate annealing temperature for 30 seconds and 72 °C for 30 seconds. Finally elongate the PCR product at 72 °C for 10

minutes.

PCR product purification: The PCR products are checked by electrophoresis, usually use 1.5% of agarose gel containing 0.05% (v/v) ethidium bromide (E1510, Sigma). Cut the clear band and purify the PCR product according to manufacturer's instructions by using the GeneJET Gel Extraction Kit (K0692, Thermo Scientific).

Ligation (for bisulfite sequencing): Measure the concentration of the PCR product using Picodrop microliter spectrophotometry (Picopet01, BobBatty). Calculate the desired amount of PCR product depends on the length of PCR products and ligate with pJET1.4 linearized plasmid backbone according to manufacturer's instructions by using CloneJET PCR Cloning Kit (K1231, Thermo Scientific).

Transformation: 10 µl of ligation product is transformed into 50 µl of competent DH5α E.coli cells at 42 °C for 90 seconds. After recovering at 37 °C for 30 minutes, transformed E.coli were grown overnight on LB Ampicillin selected plates.

Colony PCR: Use the forward and reverse primers of pJET1.4 cloning plasmid and the positive transformed E.coli as the primers and the template for colony PCR, respectively. The following program was used for colony PCR reactions: 95 °C for 5 minutes and 28 cycles of cycling with each cycle at 95 °C for 30 seconds, 56 °C for 30 seconds and 72 °C for 1 minute. Finally elongate the PCR product at 72 °C for 10 minutes.

Clean up of the colony PCR product: Generation of the mixture containing 2 µl of Exonuclease I, 20 µl of Shrimp Alkaline Phosphatase and 78 µl of H_2O. Take 10 µl of colony PCR product and mix with 4 µl of Exonuclease I and Shrimp Alkaline Phosphatase mixture. Incubate at 37 °C for 20 minutes and then at 80 °C for 15 minutes.

DNA sequencing: Big Dye Reactions were set up in a 10 μl system including 3 μl of 1 × sequencing buffer, 2 μl of BigDye Terminator v3.1 (4337457, Life-technologies), 1 μl of sequencing primer (3.2 μM), and 4 μl of cleaned up of the colony PCR product. The following program was used for the reactions: 96 °C for 1 minute and 25 cycles of cycling with each cycle at 96 °C for 10 seconds, 50 °C for 5 seconds and 60 °C for 4 minutes. Finally, DNA sequencing was carried out in the Genepool Sanger Sequencing Service department at the University of Edinburgh.

2.2.3 Methylated DNA immunoprecipitation (MeDIP)

50 μg of genomic DNA was diluted into a 1.5 ml volume with 1 × TE buffer and sonicated by biorupter device (B01010002, Diagenode) at high frequency for 25 cycles with each cycle of 30 seconds on and off. The size of the sonicated DNA was assessed by agarose gel electrophoresis. For MeDIP, the expected size is around 300-500 base pairs. 4 μg of sonicated DNA was made into a final volume of 450 μl followed by addition of 5 μg of either anti-5meC or anti-mouse IgG antibody. The sonicated DNA fragments and the antibody were incubated for 2 hours at 4 °C. 50 μl of Sheep anti-mouse Dynabeads were added and incubated for another 2 hours at 4 °C. The beads were collected by a magnet and washed in 500 μl of 1 × IP buffer for 3 times. Finally, the beads were resuspended in 100 μl of digestion buffer with Proteinase K and incubated at 55 °C overnight. After digestion, the DNA fragments were extracted by PureLink Quick PCR Purification Kit (K310002, Invitrogen) and resuspended with 100 μl of distilled water.

2.2.4 Chromatin immuoprecipitation (ChIP)

Crosslinking: cell pellets from a confluent T175 flask were resuspended in 10 ml of warm DPBS and 1/10 volume of 10 × crosslinking buffer was added. The final concentration of formaldehyde was adjusted to 1% and the cell suspension was incubated for 5 minutes at room temperature. The crosslinking reaction was stopped by 125 mM freshly made Glycine and the cells were collected by spinning at 1 400 rpm for 6 minutes at 4 °C and washed twice by cold DPBS containing 0.1 mM PMSF. The aliquoted cells (2×10^6 cells/tube) were stored at -80 °C. Cells (2×10^6) were resuspended in 5 ml cold L1 buffer containing 0.1 mM PMSF and placed on a spinning wheel for 10 minutes at 4 °C. Cells were collected by spinning at 3 000 rpm for 10 minutes at 4 °C. Then the cells were washed with cold L2 buffer containing 0.1 mM PMSF and spun down as above. Finally, cells were resuspended with 1850 µl of cold L3 buffer containing 0.5% SDS.

Sonication: the cell suspension was sonicated by biorupter device (B01010002, Diagenode) at high frequency for 45 cycles with each cycle of 30 seconds on and off. 50 µl of the sonicated chromatin was spun at 13 000 rpm for 10 minutes and the supernatant was moved into a new tube. Equal volume of 1 × TE buffer with 1% SDS and 5 µl of proteinase K were added. Then the sample was digested overnight at 65 °C. The rest of the sonicated chromatin sample was divided into equal aliquots (100 µl) and stored at -80 °C. After digestion, the DNA was extracted by the phenol/chloroform method (see methods for genomic DNA extraction) and the size of the sonicated chromatin was assessed by agarose gel electrophoresis. For ChIP, the expected size was around 300 base pairs.

Immunoprecipitation: the appropriate volume of sonicated chromatin samples, which equals to 2.5 µg of DNA, was diluted in ChIP dilution buffer in order to reduce the

SDS concentration to 0.1%. The volumes of the different samples were adjusted to equal amounts and desired antibody was added (1 μg/ 200 μl IP reaction). Then the sonicated chromatin samples and the antibody were incubated for 2 hours at 4°C. 50 μl of Protein G Dynabeads (10004D, Lifetechnologies) beads were added and incubated together for another 2 hours at 4°C. The beads were collected by a magnet and washed by Wash buffer 1 and 2 for 4 times and 3 times, respectively. The beads was then resuspended in 100 μl 1 × TE buffer with 1 μl of RNase (29 mg/ml) and incubated at 37 °C for 15 minutes. Finally, 5 μl of Proteinase K (20 mg/ml) was added and the beads were incubated at 55 °C overnight. After digestion, the DNA fragments were extracted by PureLink Quick PCR Purification Kit (K310002, Invitrogen) and resuspended with 100 μl of distilled water.

Quantitative PCR (qPCR) detection: the enrichment of the DNA fragments after ChIP was next checked by qPCR. qPCR was carried out by LightCycler 480 (Roche) with 1 × SYBR Green I Master-mix buffer containing 2.5 μM primers and desired amount of template with the conditions of denature, annealing and elongation. For all the ChIP samples, use 3 μl of purified DNA fragments as template, whereas for the input samples, use 1 μl of purified DNA fragments as template. Finally, the qPCR results were calculated by comparative Ct method, also known as the $2^{-[delta]\ [delta]\ Ct}$ method, with different dilution factor of ChIP and input samples.

2.2.5 Quantitative reverse transcription PCR (qRT-PCR)

Total RNA extraction: Cell pellets from the confluent T75 flasks were resuspended in 1 ml of Trizol reagent (15596-026, Lifetechnologies) and 200 μl of chloroform added. Samples were mixed and spun at 13 000 rpm for 15 minutes at 4 °C. The top aqueous phase was removed and precipitated with an equal volume of

isopropanol for 10 minutes at room temperature. Then the precipitated samples were spun at 13 000 rpm for 15 minutes at 4 °C. RNA pellets were washed in cold 75% (v/v) ethanol and air dried for 5 minutes at room temperature. Finally, the RNA pellets were dissolved in RNAse free water with RiboLock RNase inhibitor (EO0381, Thermo Scientific). The trace amount of DNA was digested with 2 µl of DNase I (EN0521, Thermo Scientific) in a 50 µl volume at 37 °C for 20 minutes and then the DNase was inactivated at 95 °C for 3 minutes.

cDNA synthesis: 1 µl of oligo(dT) primer (0.5 µM) was added into 4 µg of total RNA in a final volume of 11 µl. Samples were incubated at 65 °C for 5 minutes and cooled immediately on ice. 8 µl of mixture solution, including 1 µl of dNTPs (10 mM), 4 µl of 5 × FS buffer; 2 µl of DTT (0.1 M) and 1 µl of RNase inhibitor (40 units/µl), was added into each sample and incubated together at 42 °C for 2 minutes. Finally, 1 µl of Superscript II Reverse Transcriptase (18064-014, Lifetechnologies) was added selectively into the positive RT reactions and the reaction was performed at 42 °C for 2 hours.

Quantitative PCR: Quantitative PCR was carried out by LightCycler 480 (Roche) with 1 × SYBR Green I Master-mix buffer containing 2.5 µM primers and desired amount of template with the conditions of denature, annealing and elongation. Finally, the qPCR results were calculated by comparative Ct method, also known as the $2^{-[\text{delta}]\ [\text{delta}]\ Ct}$ method, with different dilution factor of ChIP and input samples.

2.2.6 Nuclear protein extraction and Western blotting

Nuclear extracts: Cell pellets collected from two confluent T175 flasks were resuspended in 1 ml of cold hypotonic NE1 Buffer with proteinase inhibitor (P8340,

Sigma) and homogenized in a 2 ml Dounce homogenizer. The disrupted cell suspension was transferred into a 2 ml Eppendorf tube and spun at 3 000 rpm for 5 minutes at 4 °C in order to collect nuclei. The supernatant was removed and the nuclear pellet was resuspended in a desired volume of cold NE1 buffer. The nuclear acids were digested on ice by 3 µl of Benzonase (70746, Merck Millipore) for 1 hour and the digestion was stopped by 450 mM NaCl. Samples were placed on a spinning wheel for 1 hour at 4 °C and then spun at 13 000 rpm for 15 minutes at 4 °C. Finally, the supernatant nuclear extract layer was transferred into a new tube.

Protein quantification: Protein concentration was measured by Bicinchoninic Acid-Copper (II) Sulfate method. First, Bicinchoninic Acid Solution (Sigma, B9643) and Copper (II) Sulfate Solution (Sigma, C2284) were mixed according to a 50:1 ratio. For each sample, take 2 µl of protein and resuspend in 1 ml mixture of Bicinchoninic Acid Solution and Copper (II) Sulfate Solution. Heat and shake the samples at 65 °C at 800 rpm for 15 minutes and place the sample on ice for 5 minutes. Measure the OD value at 562 nm wavelength by Scanning Spectrophotometer (3000, CECIL Instruments) and calculate the protein concentration according to 0.04 absorbance units are approximately equivalent to 1 µg of protein. Quantified protein extracts were stored at -80 °C.

SDS-PAGE electrophoresis: 4% or 15% separating gel was prepared and poured prior to the stacking gel. Protein was resuspended in a 1 × sample buffer system containing 0.1 M DTT and denatured at 95 °C for 5 minutes. Different sizes of the protein were separated in a suitable concentration SDS-PAGE gel at 270 V for 75 minutes.

Nitrocellulose membrane transfer: Protein separated in the SDS gel was transferred onto Nitrocellulose membrane (Bio-Rad) for 1 hour in 1 × cold transfer buffer

system at 270 V.

Western blotting: Membrane was blocked in 1 × PBS with 4% milk and 0.1% Tween 20 for 1 hour at RT. Then the membrane was probed overnight at 4 °C in the freshly made blocking solution containing the desired concentration of primary antibody concentration (see materials for antibody concentrations). Membrane was washed three times with 1 × PBS with 0.1% Tween 20 and then blocked again for 1 hour as mentioned above. Secondary antibody was added to a desired concentration (see materials for antibody concentrations) and the membrane was incubated for another 2 hours. Finally, the membrane was washed as before to remove the unspecific binding of the secondary antibody and scanned by Odyssey 3.0 Scanner (Licor biosciences).

2.2.7 Mouse ES cells culture

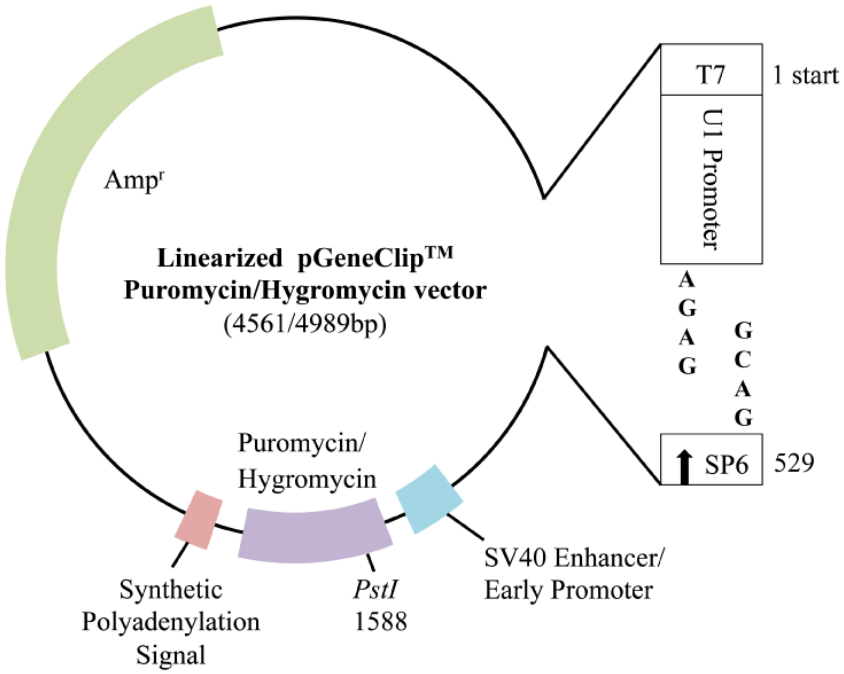
The ES cells were cultured at 37 °C and 5% CO2 in the flasks or dishes coated with 0.5% sterile gelatin in Minimum Essential Medium (MEM, 11430-030, Lifetechnologies) supplemented with 10% (v/v) Fetal Bovine Serum (FBS), 1% (v/v) sodium pyruvate, 1% (v/v) non-essential amino acids, 1% (v/v) Penicillin-Streptomycin-Glutamine (PSG), 0.05% (v/v) of 2-mercaptoethanol and desired amount of Leukemia inhibitory factor (LIF, usually 2 µl of LIF in 1 ml medium). Confluent cells were trypsinized with 0.05% warm trypsin (R001100, Lifetechnologies) and the cells were either be passaged or frozen in medium containing 10% (v/v) DMSO at -80 °C.

2.2.8 Knockdown of G9a by short hairpin RNA (shRNA)

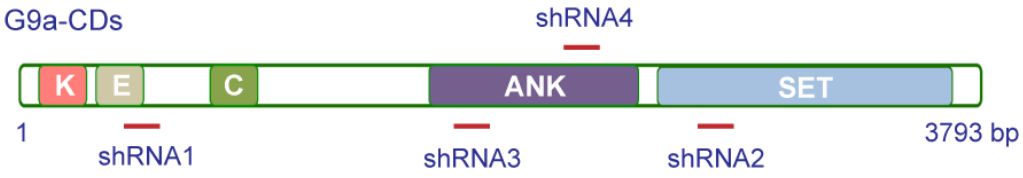
One control and four G9a cDNA specific targeted short hairpin RNA (shRNA) expression plasmids are designed and synthesized from Creative Biogene (Figure 2.1). 500 µg plasmids were linearized by ScaI (R0122S, NEB) and precipitated by ethanol.

Confluent growing ES cells were trypsinized and 5×10^6 cells were resuspended in DPBS containing 50 μg of linearized plasmid DNA in a Bio-Rad Transfection cuvette in a final volume of 800 μl. The cell suspension was gently mixed and immediately transfected by MicroPulser Electroporator (165-2100, Bio-Rad) at 800 V for 0.04 ms with a capacitance of 3 μF. After transfection, the cells were kept at room temperature for 15 minutes and seeded into 10 cm gelatinized tissue culture dishes containing the 10 ml of ES cell culture medium. The medium was replaced by fresh medium containing the desired amount of antibiotics 24 hours after transfection. The positive cells were kept growing with antibiotics until the colonies appeared.

A



B



C

	Target positions	Sequences of target positions
shRNA1	273-93	CGAGAGAGTTCATAGCTCTTT
shRNA2	3439-59	GGCGAGGTTTACTGCATTGAT
shRNA3	2410-30	GGCTCCACCTGTCTACATCAT
shRNA4	2937-57	AGCCCGAGGCTATGAGAATGT

Figure 2.1 Structure of *G9a* knockdown short hairpin RNA (shRNA) expression vector and target positions of these shRNAs within G9a.

(A) Structure of *G9a* knockdown short hairpin RNA (shRNA) expression vector (Suresliencing shRNA vector). Vectors were bought from QIAGEN company. (B) The schematic view of G9a coding sequences (G9a-CDs). Red lines represent the target positions of four *G9a* knockdown shRNAs within G9a. ANK: sequences coding for ankyrin repeats domain; SET: sequences coding for SET domain; K, E and C:sequences coding for lysince rich region, glutamic acid rich region and cysteine rich region. (C) Target positions and sequences of these shRNAs within G9a.

2.2.9 Generation of plasmids expressing mutant forms of G9a

Design two pairs of primers, including G9a-BglII-F, G9a-shMUT-R, G9a-shMUT-F, G9a-BglII-R (see 2.1.3 Primers) and make sure G9a-shMUT-R and G9a-shMUT-F match with each other and contain seven nucleotide acid mutations, which involved in the short hairpin RNA targeted regions but without changing the coding G9a amino acid sequences. Using pCAG-G9a expressing plasmid as template, first amplify two fragments of DNA using the first (G9a-BglII-F and G9a-shMUT-R) and second (G9a-shMUT-F and G9a-BglII-R) pairs of primers individually.

Purify the PCR products and perform the second round PCR using the first round two PCR products as template and G9a-BglII-F and G9a-BglII-R as primers. Purify the second round PCR product and cut by BglII. Purify the BglII cut PCR product and ligate with the pCAG-G9a expressing plasmid backbone, which is also cut by BglII. Transform the ligation product into competent E.coli cells and select the positive clones on LB ampicillin resistant plates. Pick and grow the positive bacteria clones and extract plasmids from them. Sequencing the plasmids and ensure the positive plasmid containing seven nucleotides acid mutations, which against the short hairpin RNA (shRNA) without altering the sequences of amino acid in G9a protein.

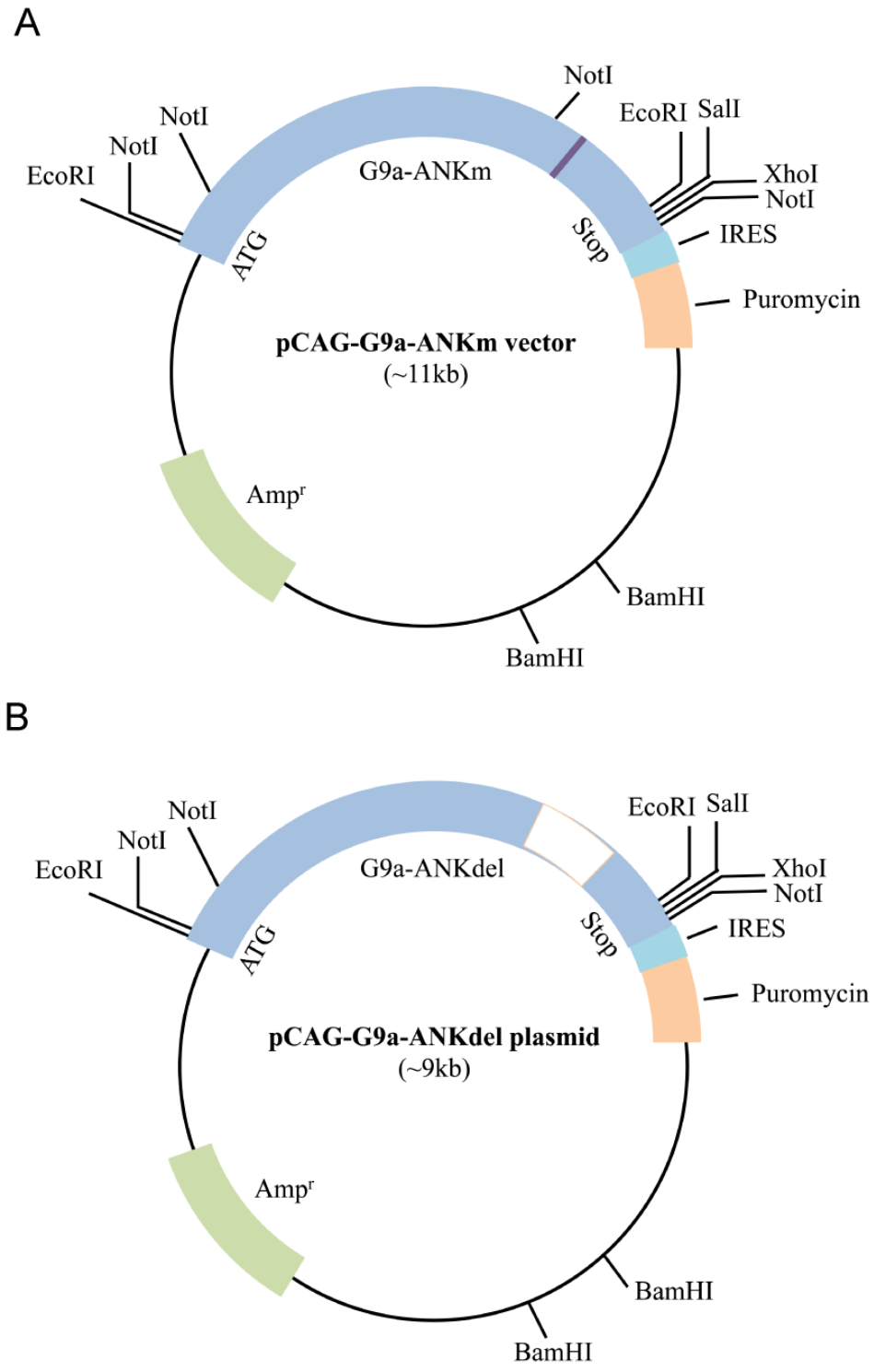


Figure 2.2 Maps of ANKm and ANKdel forms of G9a expression plasmids.

(A) The structure of ANKm G9a expression plasmid. The purple bar represents for nucleotides mutations associated with W844A and E847A within G9a protein. (B) The structure of ANKdel G9a expression plasmid. The missing section within plasmid represents for deletion of nucleotides sequences that coding for ANK domain within G9a protein. Entire coding sequences of ANKm and ANKdel G9a are shown in the appendix.

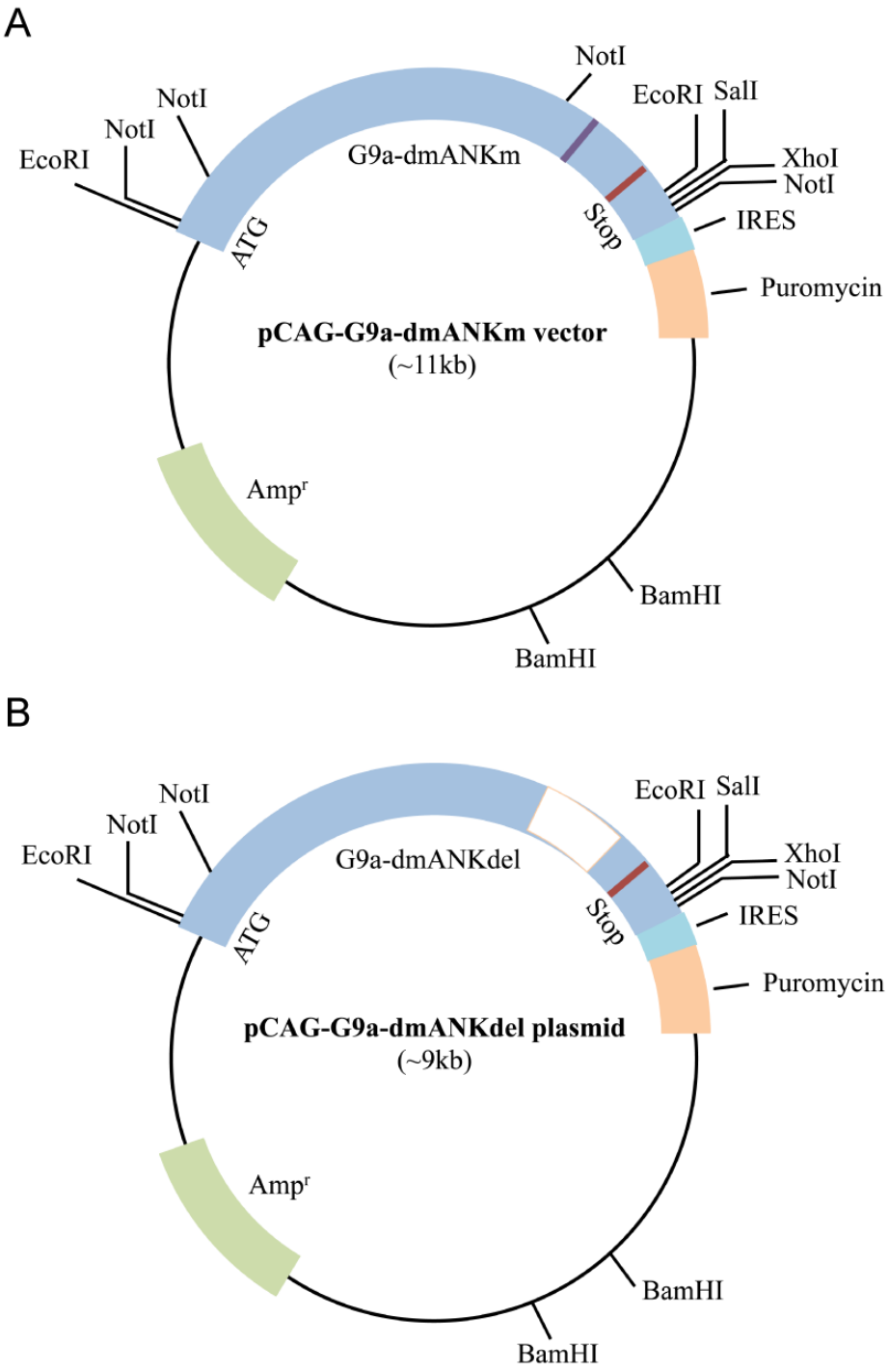


Figure 2.3 Maps of dmANKm and dmANKdel forms of G9a expression plasmids.

(A) The structure of dmANKm G9a expression plasmid. The purple bar represents for nucleotides mutations associated with W844A and E847A within G9a protein. The red bar represents for nucleotides mutations associated with NH to LE (1165-6) within G9a protein. (B) The structure of dmANKdel G9a expression plasmid. The missing section within plasmid represents for deletion of nucleotides sequences that coding for ANK domain within G9a protein. The red bar represents for nucleotides mutations associated with NH to LE (1165-6) within G9a protein. Entire coding sequences of dmANKm and dmANKdel G9a are shown in the appendix.

Next, using pCAG-G9a (resistant to shRNA) plasmid as template, generate of the ANK domain mutant (ANKm, W844A, E847A) and ANK domain deleted (ANKdel, Δ692-943aa) forms of G9a expressing plasmids by the similar strategy (Primers see 2.1.3) (Figure 2.2). Furthermore, by individually create the NH to LE (1165-6) mutation in ANKm G9a and ANKdel G9a expressing plasmids, the dmANKm G9a and dmANKdel G9a expressing plasmids are generated (Primers see 2.1.3) (Figure 2.3). Finally, all the nucleotide mutations in the plasmids that mentioned above are confirmed by DNA sequencing.

2.2.10 Generation of stable ES cell lines expressing mutant forms of G9a

First, linearize the mutant forms of G9a expressing plasmids by BamHI (R0136, NEB). Purify the linearized plasmids and transfected into the wild-type ES cells using MicroPulser Electroporator (165-2100, Bio-Rad) (Details see 2.1.8). 24 hours after transfection, select the positive cells by 1 μg/ml puromycin. Pick and grow the positive cells and perform Western blotting in order to check the exogenous G9a expression levels.

Choose the stable and highest exogenous G9a expression cells and transfected by the linearized G9a short hairpin RNA expressing plasmid. 24 hours after transfection, select the positive cells by 200 μg/ml hygromycin. Pick and grow the positive cells and perform Western blotting and qPCR in order to check the endogenous G9a expression levels. Choose the stable and lowest endogenous G9a expression cells for the downstream experiments.

2.2.11 Mouse embryos handling

Pregnant mice were killed according to the manipulation instructions and the uterus with the embryos was collected. Embryos were dissected from the uterus without impairment of amniotic membrane and maternally derived blood was washed away by DPBS 5 times. The amniotic membrane was punctured and the embryos were dissected in individual dishes containing DPBS. Then the extraembryonic tissues were removed from the embryos. Finally, the embryos were washed 3 times by DPBS and used for downstream experiments.

Chapter 3 Loss of imprinted DNA methylation in G9a and GLP null ES cells

3.1 Introduction

In mammals, among identified histone lysine methyltransferases (HKMTs), G9a and GLP serve as the main modifiers for the establishment and maintenance of mono- and di-methylation at lysine 9 on the tail of histone H3 (H3K9me1 and H3K9me2). Genome-wide analysis of histone modifications demonstrated that H3K9me2 covers nearly half of the genome and both H3K9me1 and H3K9me2 are found to be associated with transcriptional repression (Lienert et al., 2011). In human cells, the large lamina associated chromatin domains, which are normally silenced, contain high levels of H3K9me2 (Liu et al., 2011). In mouse ES cells, the majority of G9a repressed genes are localized to H3K9me2 marked nuclear periphery and these genes can be re-activated by G9a inactivation (Kind et al., 2013).

Actually, it is found that global distribution of H3K9me2 in pluripotent and differentiated ES cells are quite similar (Filion and van Steensel, 2010). In addition, both H3K9me2 and H3K9me3 cover pericentric and telomeric regions in many species, where they function in suppressing recombination and silencing the transposable elements (Bickmore and van Steensel, 2013). Furthermore, heterochromatin mark H3K9me2 not only plays a role in early developmental stages, but also in injury and tumourgenesis (Yu et al., 2013; Li et al., 2014). One of the best examples is H3K9me2 appears functionally during epithelial-mesenchymal transition (EMT). In Claudin-low breast cancer (CLBC) cells, it has been found that G9a is responsible for DNA methylation that is present at promoter of E-cadherin, resulting decreased expression of E-cadherin, a hallmark of EMT (Dong et al., 2012). In particular, H3K9me2 is also associated with repression of

tumor suppressor genes, since it is erased from promoters of reactivated tumor suppressor genes in cancer cells that have been treated with inhibitors of either DNMTs or HDACs (McGarvey et al., 2006).

In fungi and plants, H3K9 methylation often associates with DNA methylation (Tamaru and Selker, 2001). In mammals, H3K9 di- and tri-methylation also play a role in the maintenance of DNA methylation. It has been found that knockout of *G9a* leads to a 50% loss of DNA methylation in ES cells, especially at repetitive elements, such as retrotransposons and major satellite repeats (Dong et al., 2008; Leung et al., 2011). Apart from repetitive regions, histone methyltransferase G9a also plays a role in *de novo* and maintenance of DNA methylation at protein coding regions. For instance, during differentiation of *G9a-/-* ES cells, promoters of some pluripotent genes, such as *Oct4* and *Nanog*, fail to gain DNA methylation (Athanasiadou et al., 2010; Epsztejn-Litman et al., 2008; Feldman et al., 2006). In addition, at some imprinted control regions (ICRs), maintenance of imprinted DNA methylation also requires presence of G9a, since DNA methylation that is present at Prader-Willi syndrome ICR is reduced and the associated imprinted gene, *Snrpn*, is bi-allelically expressed in *G9a-/-* ES cells (Xin et al., 2003). Interestingly, loss of DNA methylation in *G9a* deficient ES cells can be partially restored by catalytically inactive G9a (Dong et al., 2008; Myant et al., 2011; Zhao and Epstein, 2008). Taken together, these observations indicate that apart from depositing H3K9me1 and H3K9me2, histone methyltransferase G9a also affect DNA methylation and the role of G9a in either the maintenance or establishment of DNA methylation is independent of its catalytic activity.

A previous study in the lab combining methylated DNA affinity purification with microarray analysis suggested that some maternal ICRs are hypomethylated in *G9a-/-* ES

cells (Ausma Termanis, unpubished). Depending on this observation, the first aim of this chapter was to validate the data from this genome wide study. Second, since only maternal, but not paternal, ICRs are located at promoters of imprinted genes, I set out to investigate whether paternal ICRs also lose DNA methylation in *G9a-/-* ES cells. In addition, since catalytically inactive G9a can partially restore loss of DNA methylation in *G9a* deficient ES cells, I also set out to investigate whether loss of imprinted DNA methylation can be restored by re-expression of either wild-type or catalytically inactive G9a in *G9a-/-* ES cells. Finally, I was also interested in investigating the patterns of imprinted DNA methylation in *Glp*-/- ES cells, since G9a/GLP complex is predominant form of either G9a or GLP exist *in vivo* (Ueda et al., 2006).

3.2 Results

3.2.1 DNA methylation is lost from ICRs in *G9a-/-* ES cells

Validation of promoter microarray data by bisulfite DNA sequencing

I carried out bisulfite DNA sequencing at *Ankrd50* promoter, which was identified to be highly methylated in both wild-type and *G9a-/-* ES cells in the genome wide study of affinity purification of methylated DNA coupled with hybridization to promoter microarrays. I found that *Ankrd50* promoter is highly methylated in both wild-type and *G9a-/-* ES cells (Figure 3.1A), although compared with wild-type ES cells, there is a subtle reduction of DNA methylation at *Ankrd50* promoter in *G9a-/-* ES cells. The maternal imprinted gene, *Mest*, also known as *Peg1*, contains a DMR within its promoter. In previous genome-wide study, compared with wild-type ES cells, *Mest* was one of imprinted genes, which showed hypo-methylation in *G9a-/-* ES cells. *Igf2r* is another maternal imprinted gene, which contains two DMRs. The first DMR (DMR1) is located

at its promoter, which plays a role in the regulation of mono-allelic expression of *Igf2r*. The second DMR (DMR2) is located within *Igf2r* gene body region and overlaps with promoter of *Airn*, a gene encodes a long non-coding RNA. DMR2 has been shown to be associated with the regulation of mono-allelic expression of *Airn*. In order to validate the previous genome-wide DNA methylation study, I performed bisulfite DNA sequencing at promoter of *Mest* as well as DMR2 of *Igf2r*. Compared with wild-type ES cells, I found that these two regions displayed hypo-methylation in *G9a-/-* ES cells, confirming requirement of G9a for the protection of DNA methylation at these maternal ICRs (Figure 3.1A).

Paternal ICRs also display hypo-methylation in *G9a-/-* ES cells

In mammals, paternal ICRs are located at the intergenic regions of imprinted loci. Since the microarray only spanned promoters of protein coding genes, conclusions could not be made about the patterns of imprinted DNA methylation at intergenic ICRs in *G9a-/-* ES cells. In order to investigate imprinted DNA methylation patterns at paternal ICRs in *G9a-/-* ES cells, I performed bisulfite DNA sequencing at *Igf2-H19* ICR. Compared with wild-type ES cells, *Igf2-H19* ICR is hypo-methylated in *G9a-/-* ES cells, which indicates that not only maternal, but also some paternal ICRs lack imprinted DNA methylation in *G9a-/-* ES cells (Figure 3.1B).

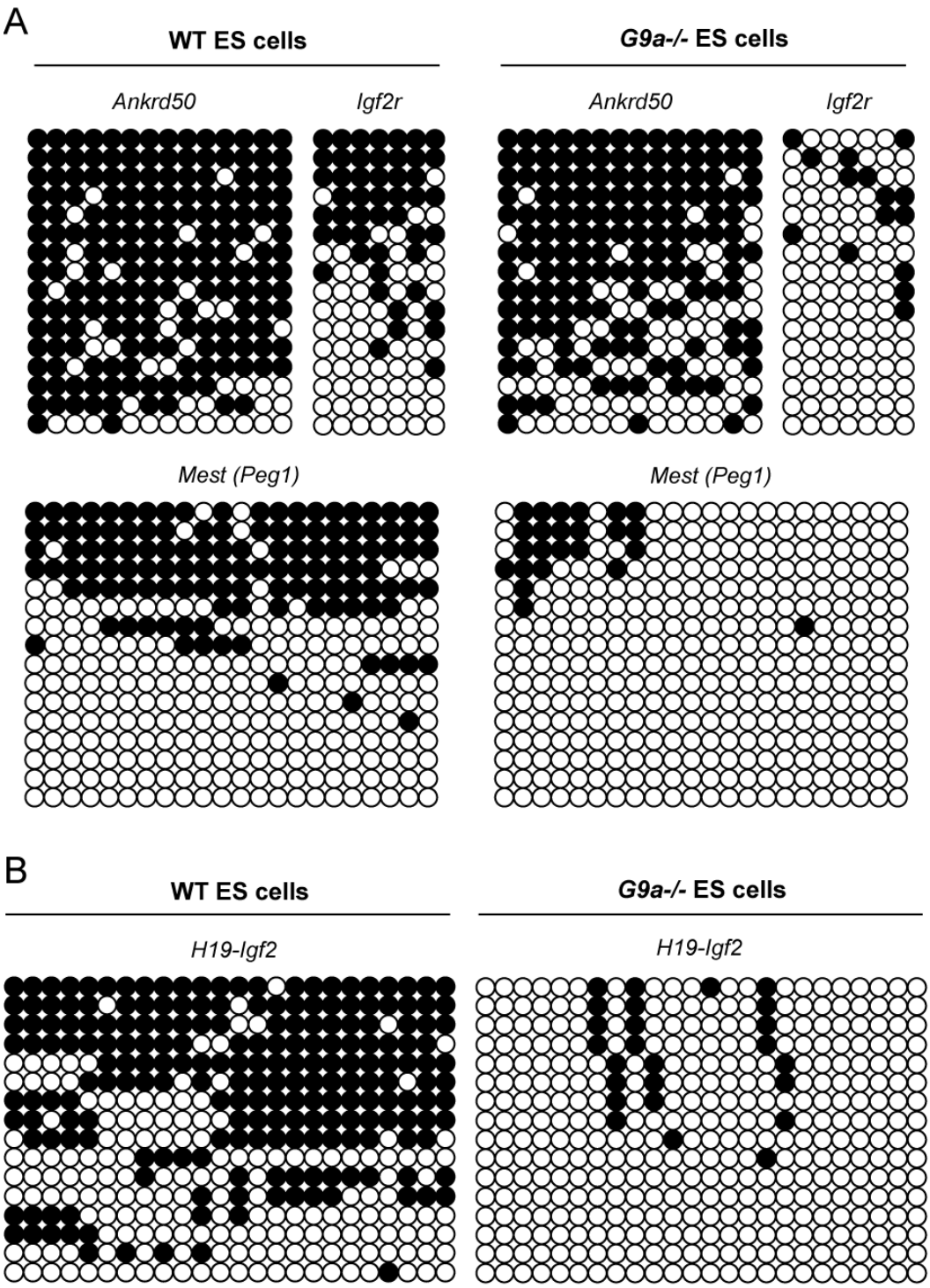


Figure 3.1 DNA methylation is lost at the ICRs in *G9a-/-* ES cells.

(**A**) Bisulfite DNA sequencing at maternally methylated ICRs, *Igf2r* and *Mest,* in wild-type and *G9a-/-* ES cells. *Ankrd50* serves as a positive control region, which is methylated according to the microarray data in both the wild-type and *G9a-/-* ES cells. (**B**) Bisulfite DNA sequencing at paternally methylated ICR, *H19-Igf2*, in wild-type and *G9a-/-* ES cells. The black circles represent the methylated CpG sites, whereas the white circles indicate the unmethylated CpG sites. Each row represents a single DNA strand.

3.2.2 G9a is required for the DNA methylation at a large number of ICRs

In order to validate bisulfite DNA sequencing results and investigate imprinted DNA methylation patterns at more imprinted loci, I carried out 5meC methylated DNA immunoprecipitation (MeDIP) on wild-type and *G9a-/-* ES cells. Since *Ndufa* and *Ankrd50* are two non-imprinted regions that display low and high levels of DNA methylation, respectively. In wild-type and *G9a-/-* ES cells, according to the promoter microarray data, these two regions served as negative and positive control regions for MeDIP experiments.

5meC MeDIP results showed that *Ndufa* and *Ankrd50* regions display low and high levels of DNA methylation as expected and these levels were comparable between wild-type and *G9a-/-* ES cells (Figure 3.2). Apart from these control regions, I also carried out MeDIP followed by qPCR at some ICRs, including five maternal ICRs, including *Igf2r*, *Snrpn*, *Zac1*, *Peg3* and *Peg10* and one paternal ICR *Igf2-H19*. Compared with wild-type ES cells, 5meC enrichment at all the tested ICRs displays significant reduction in *G9a-/-* ES cells, indicating a widespread loss of imprinted DNA methylation in *G9a-/-* ES cells. Therefore it could be deduced that G9a deficiency affects patterns of DNA methylation at all the imprinted loci in genome.

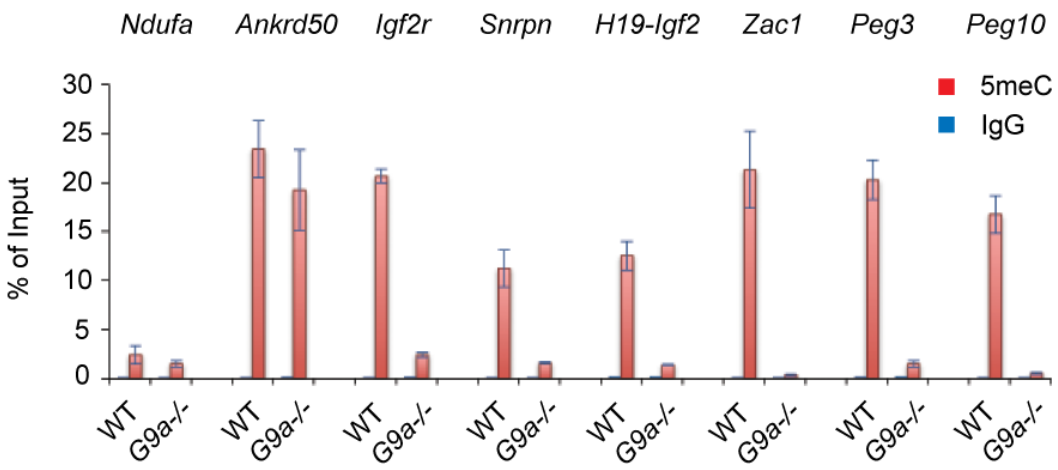


Figure 3.2 Imprinted DNA methylation is absent from all the ICRs in *G9a-/-* ES cells.

MeDIP detecting loss of imprinted DNA methylation from paternally (*H19-Igf2*) and maternally (*Igf2r*, *Snrpn*, *Zac1* etc.) methylated ICRs in *G9a-/-* ES cells. *Ankrd50* and *Ndufa* serve as the positive (methylated) and negative (unmethylated) control regions, respectively. The Y-axis value represents percentage of enrichment normalized to input DNA. IgG (blue bars) was employed as the negative control for 5meC MeDIP. Error bars represent 2 times standard deviations of technical triplicates.

3.2.3 Imprinted DNA methylation is lost in the *Glp-/-* ES cells

Since G9a/GLP complex is the predominant form of either G9a or GLP *in vivo* and GLP also plays a role in the establishment of H3K9 methylation, I set out to investigate the patterns of imprinted DNA methylation in *Glp-/-* ES cells. Similar to in *G9a-/-* ES cells, bisulfite DNA sequencing results showed that *Ankrd50* promoter was highly methylated in *Glp-/-* ES cells (Figure 3.3). However, both the ICRs that I examined, including *Igf2r* DMR2 and promoter of *Snrpn*, displayed hypo-methylated in *Glp-/-* ES cells (Figure 3.3). This result indicates that both the histone methyltransferases, G9a and GLP, are essential for the normal patterns of imprinted DNA methylation in ES cells.

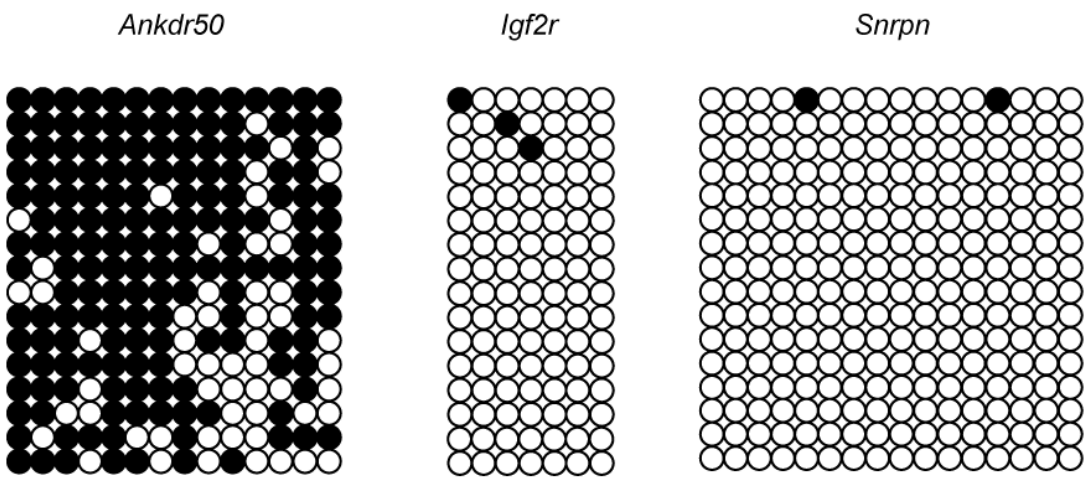


Figure 3.3 Imprinted DNA methylation is absent from the ICRs in *Glp-/-* ES cells.

Bisulfite DNA sequencing at the control *Ankrd50* promoter and the ICRs in *Glp-/-* ES cells. Compared with *Ankrd50* locus, the majority of imprinted DNA methylation at *Igf2r* and *Snrpn* ICRs was lost. Black circles indicate the methylated CpG sites, whereas white circles represent the unmethylated CpG sites. Each row represents a single DNA strand.

3.2.4 Neither wild-type nor catalytically inactive G9a can restore the loss of imprinted DNA methylation in the *G9a-/-* ES cells

It was previously shown that loss of DNA methylation could be partially restored by re-expression of either wild-type or catalytically inactive G9a in *G9a-/-* ES cells (Dong et al., 2008). In order to investigate whether loss of imprinted DNA methylation in *G9a-/-* ES cells can be restored, I used *G9a-/-* ES cell lines that re-expressing either Flag tagged transgene wild-type G9a (*G9a-/-* TgG9a) or mutant transgene G9a (*G9a-/-* TgG9a^{F1205Y}) that carrying a single point mutation substituting phenylalanine 1205 with tyrosine. To determine G9a protein and H3K9 methylation levels, I performed Western blotting assays and found that wild-type ES cells express five isoforms of G9a protein, among which two are highly expressed (Figure 3.4A). Whereas *G9a-/-* TgG9a or *G9a-/-* TgG9a^{F1205Y} ES cells expresses only one isoform of either Flag tagged wild-type G9a-short protein or catalytic inactive G9a-short protein, respectively, and both of these exogenous forms of G9a-short protein were stably expressed in *G9a-/-* ES cells (Figure 3.4A). In agreement with published data, H3K9me2 and H3K9me3 were restored in *G9a-/-* ES cells that expressing wild-type G9a and the levels are comparable to wild-type ES cells. However, the levels of H3K9me2 and H3K9me3 were not restored in *G9a-/-* ES cells that expressing catalytically inactive G9a (Figure 3.4A).

I then investigated the patterns of imprinted DNA methylation in these rescued *G9a-/-* ES cells. I carried out bisulfite DNA sequencing at four ICRs, including *Igf2r* DMR2, *IG*-DMR and promoters of *Snrpn* and *Zac1* (Figure 3.4B and C and some data not shown). I found that neither wild-type nor catalytically inactive G9a restored loss of imprinted DNA methylation at ICRs in *G9a-/-* ES cells (Figure 3.4B-C), with the

exception of *Igf2r* DMR2, which displayed partially restored DNA methylation in *G9a-/-* TgG9a cell line.

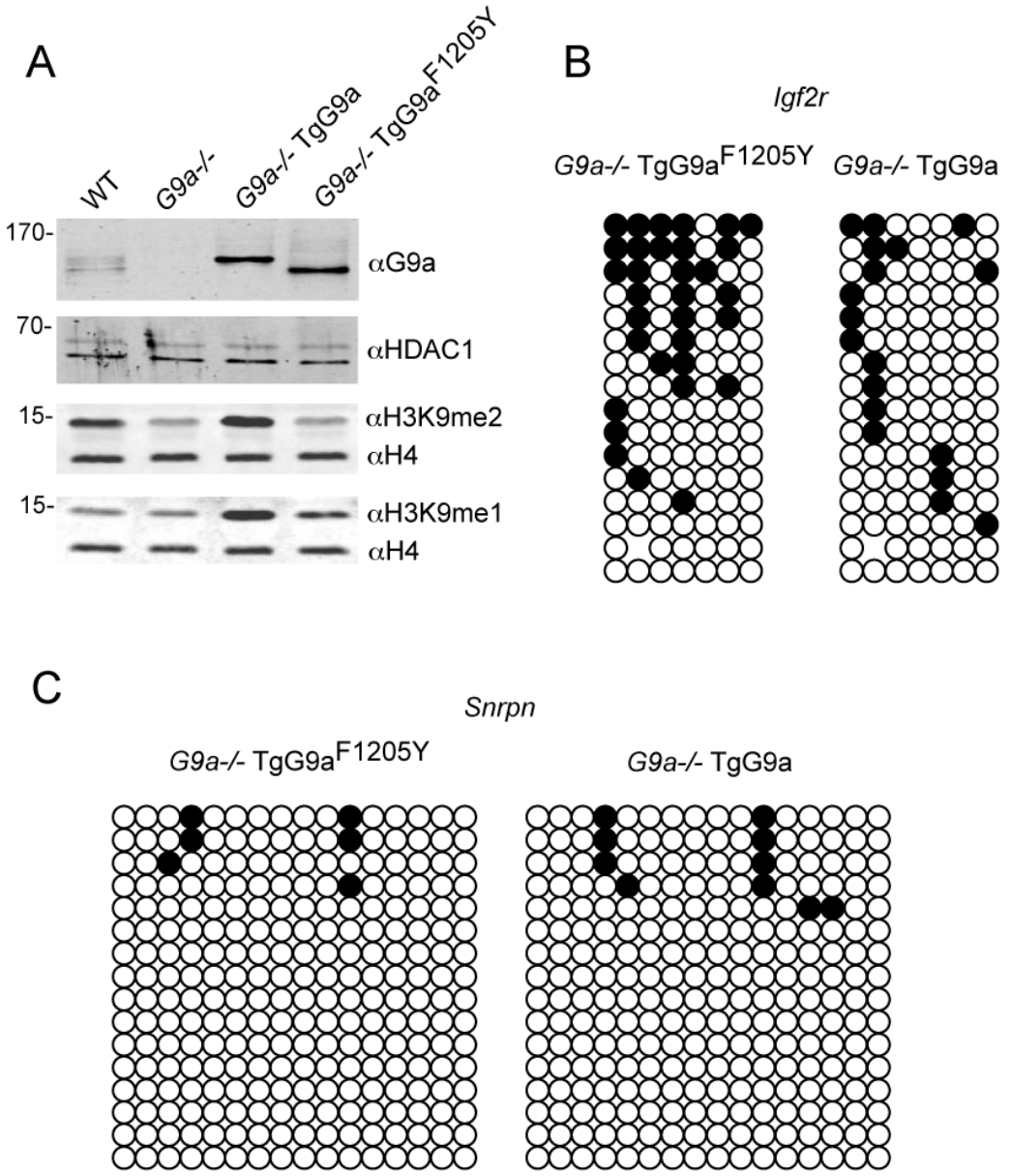


Figure 3.4 Re-expression of G9a can not restore the loss of imprinted DNA methylation in *G9a-/-* ES cells.

(A) Western blots detecting G9a, H3K9me2 and H3K9me1 on nuclear extracts from wild-type, *G9a-/-* and *G9a-/-* ES cells expressing either wild-type (TgG9a) or catalytically inactive form of G9a (TgG9a^{F1205Y}). **(B, C)** Bisulfite DNA sequencing at *Igf2r* DMR2 and *Snrpn* ICR in *G9a-/-* TgG9a and *G9a-/-* TgG9a^{F1205Y} ES cells. *IG-DMR* and *Zac1* ICRs display similar petterns (not shown).

In order to investigate the imprinted DNA methylation patterns at other imprinted loci, I carried out 5meC MeDIP on *G9a-/-*, *G9a-/-*TgG9a and *G9a-/-* TgG9a^{F1205Y} ES cells. Again, *Ndufa* and *Ankrd50* served as negative and positive control regions, respectively. At *Ndufa* locus, all these three ES cell lines display expected low DNA methylation levels. However, only in *G9a-/-* and *G9a-/-* TgG9a ES cell lines, *Ankrd50* locus displayed high levels of DNA methylation. Reduced DNA methylation at *Ankrd50* locus in *G9a-/-* TgG9a^{F1205Y} ES cells could be caused by long time cell culturing *in vitro* that was required for generation of these cells.

In addition to the control regions, six maternal ICRs: *Igf2r*, *Snrpn*, *Zac1*, *Peg3*, *Peg10* and *Peg13* and two paternal ICR *Igf2-H19* and *Rasfrf1* are also investigated. All of these analyzed ICRs displayed comparably low levels of DNA methylation, which did not differ significantly between *G9a-/-* ES cells and ES cells that expressing either TgG9a or TgG9a^{F1205Y} (Figure 3.5). Taken together, these results are consistent with previous studies (Chotalia et al., 2009), showing that unlike non-imprinted DNA methylation, once imprinted DNA methylation is lost, it cannot be restored. This also applies in the case of *G9a-/-* ES cells.

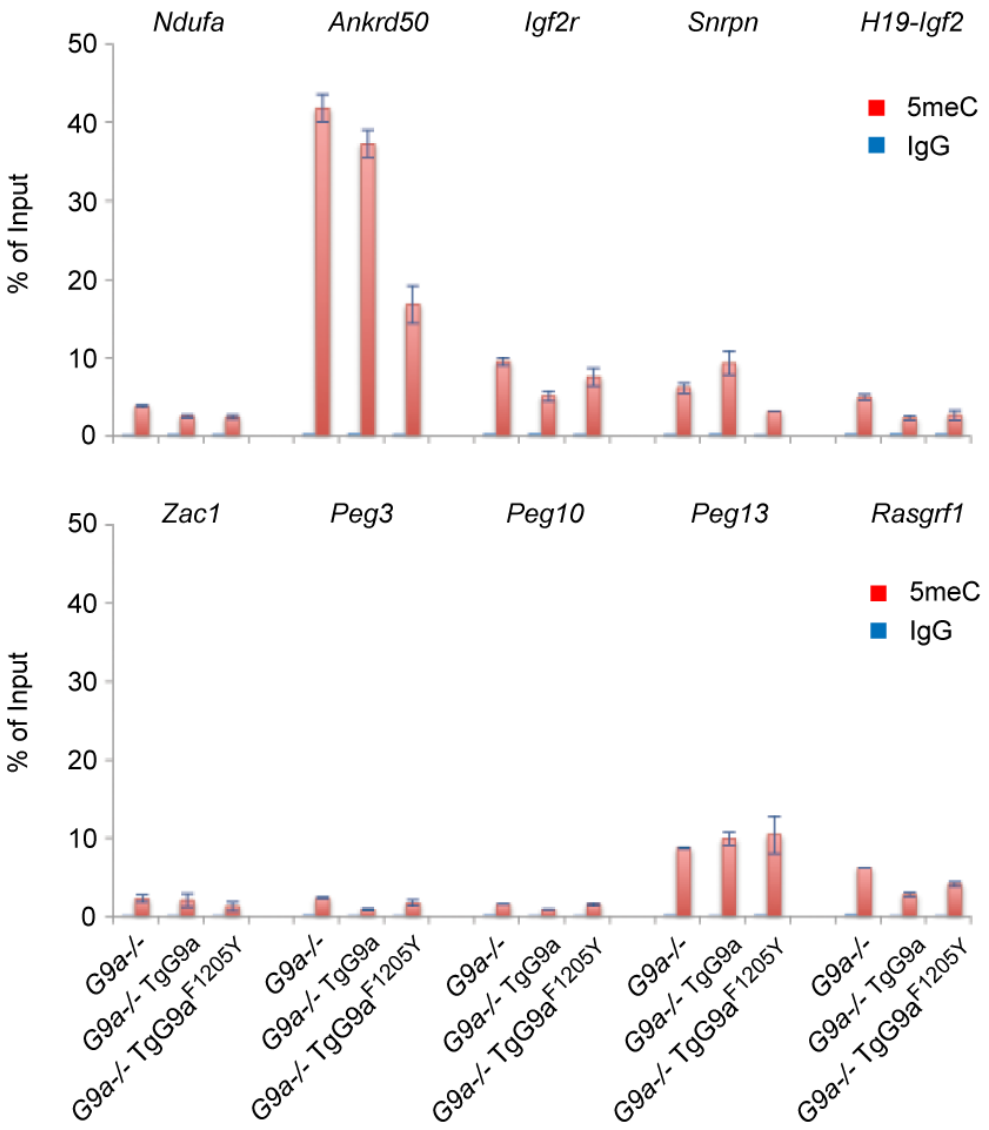


Figure 3.6 Levels of imprinted DNA methylation in *G9a-/-*, *G9a-/-* TgG9a and *G9a-/-* TgG9a^{F1205Y} ES cells.

Analyses of imprinted DNA methylation by MeDIP indicates that in the *G9a-/-* TgG9a and *G9a-/-* TgG9a^{F1205Y} ES cells, the methylation level at the ICRs remains low. *H19-Igf2* and *Rasgrf1* are two paternally methylated imprinted control regions (ICRs) and all the other ICRs are maternally methylated imprinted loci. *Ndufa* and *Ankrd50* serve as the negative and positive control regions, respectively. Y-axis value represents percentage of enrichment normalized to input DNA. Mouse IgG (blue bars) is a negative control for the anti-5meC antibody. Error bars represent 2 times standard deviations of technical triplicates. *G9a-/-* TgG9a represents *G9a-/-* ES cells expressing wild-type G9a. *G9a-/-* TgG9a^{F1205Y} represents *G9a-/-* ES cells expressing catalytically inactive G9a.

3.3 Discussion

By carrying out methylated DNA affinity purification coupled with hybridization to promoter microarrays, previous PhD student Ausma Termanis found that compared with wild-type ES cells, several maternal ICRs displayed hypo-methylation in *G9a-/-* ES cells (Figure 1.9A and B). In order to validate this observation, I performed bisulfite DNA sequencing at maternal ICRs, including *Igf2r* and *Mest* (*Peg1*) (Figure 3.1A). Together with the data at *Ankrd50* locus, which displays no difference of DNA methylation in *G9a-/-* ES cells, I found that maternal imprinted DNA methylation is lost in *G9a*-/- ES cells. In addition, I found that paternal ICRs also underwent DNA demethylation in the absence of G9a in ES cells (Figure 3.1B). Furthermore, 5meC MeDIP allowed me to investigate the DNA methylation patterns at a larger group of ICRs and consistently with bisulfite sequencing results, I found that all the investigated ICRs displayed loss of DNA methylation in *G9a*-/- ES cells (Figure 3.2).

I also carried out bisulfite DNA sequencing in *Glp*-/- ES cells and found that the levels of imprinted DNA methylation are also impaired (Figure 3.3). The possible explanation is that G9a/GLP complex serves as the predominant form of either G9a or GLP in ES cells and disruption of either G9a or GLP leads to instability of the other partner protein in heterodimeric complex (Tachibana et al., 2005; Ueda et al., 2006). Consistently, it has also been shown that G9a cannot compensate for the loss of GLP function *in vivo* and *vice versa* (Tachibana et al., 2005). Therefore, loss of imprinted DNA methylation in *G9a-/-* and *Glp-/-* ES cells indicates that both the histone methyltransferases, G9a and GLP, might play a role in the protection of imprinted DNA methylation in ES cells.

Actually, previous studies have demonstrated that histone methyltransferases G9a and GLP are required for DNA methylation in ES cells, since reduced DNA methylation at major satellite and unique loci, including *Wfdc15* and *Mega2a* was found in G9a and GLP null cells (Dong et al., 2008; Epsztejn-Litman et al., 2008; Tachibana et al., 2008). Moreover, it was shown that G9a is essential for the maintenance of imprinted DNA methylation at ICR of *Snrpn* (Xin et al., 2003). I have expanded the previous understandings and demonstrated a more widespread role of G9a in the protection of imprinted DNA methylation in ES cells. As *G9a*-/- ES cells were generated by gene conversions (Tachibana et al, 2002) and the CpG sites are heavily methylated at G9a target promoters in wild-type ES cells, which is used to generate *G9a-/-* ES cells (Tachibana et al., 2008), these wild-type ES cells should contain the normal patterns of imprinted DNA methylation at ICRs originally. Since compared with wild-type ES cells, imprinted DNA methylation is lost in *G9a* and *Glp* knockout ES cells, it is likely that G9a and GLP are required for the maintenance of imprinted DNA methylation in ES cells.

However, a question arises that by what mechanism G9a may facilitate the maintenance of imprinted DNA methylation in ES cells. Since G9a is essential for the establishment of H3K9me2 and STELLA, an H3K9me2 binding protein, is required for the protection of maternally inherited DNA methylation (including the imprinted DNA methylation) and imprinted DNA methylation present on the paternally derived chromosomes in zygote (Nakamura et al., 2007; Nakamura et al., 2012; Tachibana et al., 2005; Ueda et al., 2006), it is possible that G9a maintains the imprinted DNA methylation by depositing H3K9me2 at ICRs. On the other hand, since G9a protein *per se* could interact with DNMTs via its N-terminal domain or its ANK domain (Bestor et

al., 1988; Fatemi et al., 2002; Gowher et al., 2005), another possibility is that G9a maintains imprinted DNA methylation by directly recruiting DNMTs to ICRs in ES cells.

Since WIZ, a C2H2 zinc finger protein, is also involved in G9a/GLP complex and knockdown of *Wiz* leads to reduction of G9a and GLP (Tachibana et al., 2005; Ueda et al., 2006), it is possible that *Wiz* deficient cells would also display loss of imprinting. It is yet unknown whether WIZ can function as a potential DNA binding protein *in vivo* and which DNA motifs it may preferentially bind to. Therefore, it is not clear, whether ICRs may contain sequence specific motifs, which WIZ may bind to and recruit G9a/GP complex.

Previous studies have shown that by reintroduction of a wild-type G9a, approximately 50% of the reduced DNA methylation in *G9a-/-* ES cells can be restored and the catalytic activity of G9a is not required for the *de novo* DNA methylation (Dong et al., 2008). However, in my study, I found that loss of imprinted DNA methylation fail to be restored in either wild-type or catalytically inactive G9a transgene expressing *G9a-/-* cells (Figure 3.4B, C and 3.5).

There are several possibilities that may explain the failure of re-establishment of imprinted DNA methylation in *G9a-/-* ES cells. First, after imprinted DNA methylation is lost, un-methylated DNA binding protein may occupy the previously methylated ICRs and prevent DNMTs to re-establish imprinted DNA methylation. In addition, many studies have demonstrated that DNMT3A and DNMT3L play an essential role in the establishment of imprinted DNA methylation during development of gametes (Bourc'his et al., 2001; Hata et al., 2002; Kaneda et al., 2004). Although both DNMT3A and DNMT3L are expressed in ES cells as well as in gametes, it is unknown whether DNMT3A and DNMT3L form the same complex in ES cells to allow the establishment

of accurate DNA methylation at ICRs. Furthermore, it has also been shown that DNMT3L can directly bind to the amino terminus of histone H3 and this interaction can be abolished by presence of H3K4me2 or H3K4me3. Therefore, defects of re-establishment of imprinted DNA methylation in *G9a-/-* ES cells could also due to existence of distinct histone modifications at ICRs. Moreover, differences between transcriptomes in gametes and ES cells might serve as another explanation for the failure of restoring imprinted DNA methylation, since in some cases the oocyte-specific transcription across ICRs has been proposed to play a role in the establishment of imprinted DNA methylation in gametes (Chotalia et al., 2009).

Chapter 4 G9a is essential for the maintenance of imprinted DNA methylation in ES cells

4.1 Introduction

Previous studies indicated that G9a is essential for the protection of imprinted DNA methylation at *Snrpn* locus and loss of G9a leads to bi-allelic expression of *Snrpn* (Dong et al., 2008). In my previous study, I found that all the investigated ICRs displayed hypo-methylation in *G9a-/-* ES cells. Taken together, these observations suggested that G9a plays an important role in the protection of DNA methylation at imprinted loci in ES cells.

Since *G9a-/-* ES cells were derived from the wild-type ES cells by gene conversions (Tachibana et al., 2002) and imprinted DNA methylation is established during the development of gametes (Davis et al., 1999; Lucifero et al., 2002)(Figure 1.4), loss of imprinting in *G9a-/-* ES cells suggests that G9a might play a role in the maintenance of imprinted DNA methylation in ES cells. Therefore, the aim of this chapter is to generate stable *G9a* knockdown ES cell lines by expression of a short hairpin RNA (shRNA) and investigate whether G9a plays a role in the maintenance of imprinted DNA methylation in ES cells.

4.2 Results

4.2.1 Knockdown of *G9a* causes the reduction of DNA methylation and H3K9me at the ICRs

In order to identify the most efficient G9a knockdown short hairpin RNA (shRNA), I first tested a set of four different shRNA expression plasmids (information see Figure 2.1). Together with one control shRNA expression plasmid, I linearized these

plasmids by *SalI* and generated one shCont and four shG9a pool clonal ES cell lines in wild-type TT2 ES cells (see materials and method). I then performed Western blotting assays against G9a in these cell lines and found that compared with shCont ES cell line, shRNA-2 *G9a* knockdown pool ES cell line contains the lowest level of G9a among the four pool ES cell lines (Figure 4.1A). This result suggested that shRNA-2 cell line is the most efficient *G9a* knockdown pool clonal ES cell line.

However, within shRNA-2 pool clonal ES cell line, since individual positive cell may integrate with different copies of linearized plasmids, it is required to pick single clonal *G9a* knockdown ES cell line. I picked six independent colonies from shRNA-2 pool clonal ES cells and expanded these single cells for approximately two weeks *in vitro*. I then tested G9a mRNA levels by RT-qPCR and found that among the six independent single clonal shG9a cell lines, shG9a-2 displayed the lowest level of G9a mRNA when compared to control ES cell line (Figure 4.1B). In addition, the levels of G9a mRNA in wild-type and shCont ES cell lines were comparable, whereas shRNA-2 pool clonal ES cell line displayed reduction of G9a mRNA by 60%.

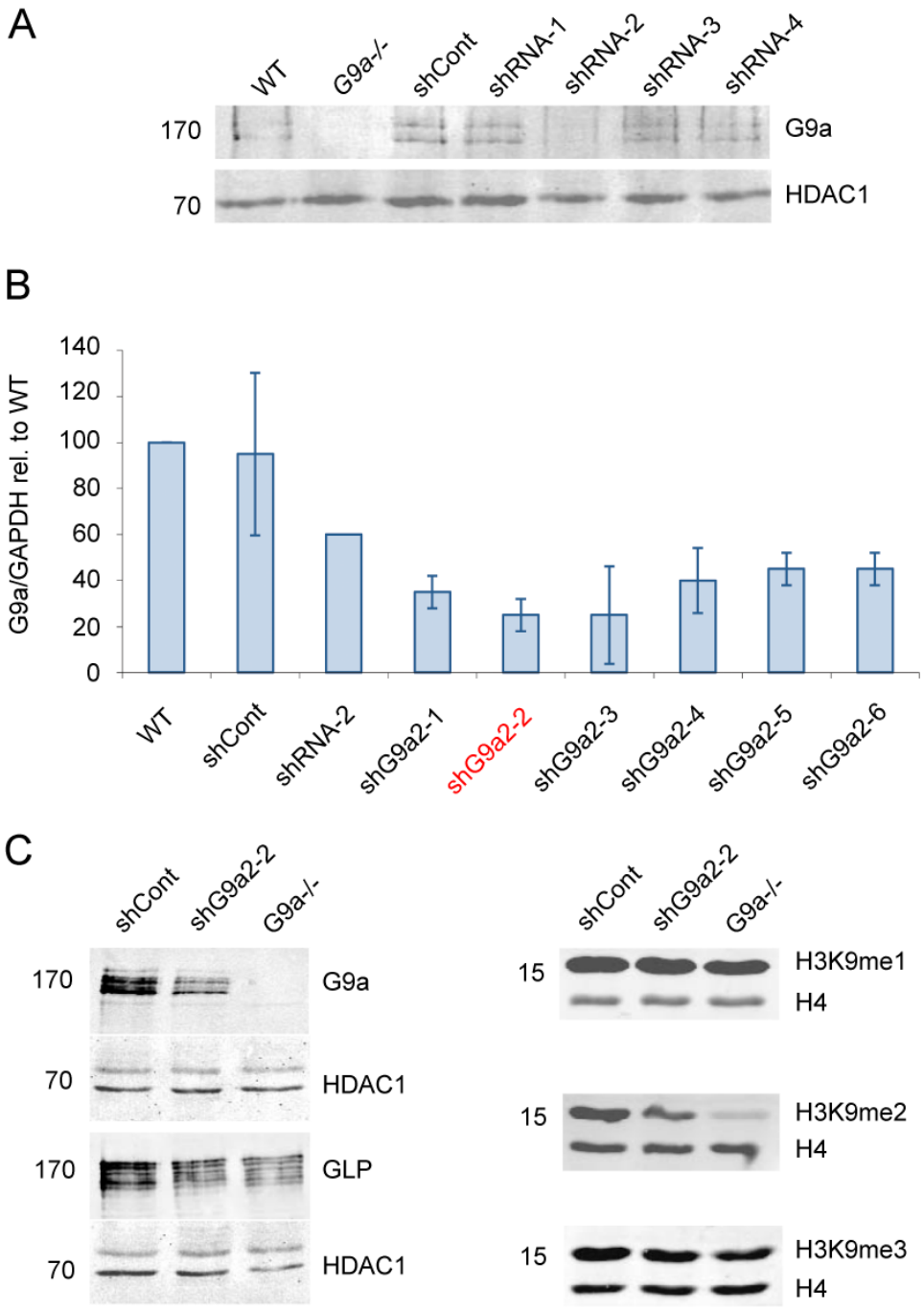


Figure 4.1 Generation of stable shG9a clonal cell line in TT2 ES cells.

(**A**) Western blots detecting G9a and HDAC1 levels in wild-type, *G9a-/-*, one control short hairpin RNA (shCont) and four G9a knockdown short hairpin RNA (shRNA) pool clonal ES cell lines. HDAC1 serves as the loading control for G9a. (**B**) Quantitive PCR detecting the mRNA levels of G9a in wild-type, *G9a-/-*, shCont pool clonal cell line, shRNA-2 pool clonal cell line and six individual G9a knockdown single clonal ES cell lines. The single clonal cell line marked in red were used for the further analyses. *Gapdh* serves as a control gene for *G9a*. (**C**) Western blots detecting G9a, GLP and H3K9me levels in shCont single clonal cell line, shG9a-2 single clonal cell line and *G9a-/-* ES cells. HDAC1 and H4 are used as the loading controls for G9a and H3K9me, respectively.

To confirm the level of G9a protein was also reduced in the shG9a-2 single clonal ES cell line, I extracted nuclear proteins and carried out Western blotting assays. The results indicated that shG9a-2 single clonal ES cells displayed a 70% reduction of G9a protein when compared with shCont ES cells (Figure 4.1C). In addition, the level of GLP, a partner protein of G9a, was also reduced in the shG9a-2 single clonal ES cells and is comparable to the level in *G9a-/-* ES cells. Since both the levels of G9a mRNA and protein were significantly reduced in shG9a-2 single clonal ES cells, shG9a-2 single clonal ES cell line has been chosen as the best single *G9a* knockdown clonal ES cell line for downstream experiments.

Previous studies have shown that G9a and GLP function as the main histone methyltransferases for the establishment and maintenance of H3K9me1/2 and G9a/GLP complex is the predominant form of either G9a or GLP that exist *in vivo* (Tachibana et al., 2002; Tachibana et al., 2005; Ueda et al., 2006), therefore, knockdown of *G9a* was likely to affect the levels of H3K9me in ES cells. To investigate the levels of H3K9me in shG9a ES cells, I carried out Western blotting assays on nuclear extracts from shCont, shG9a-2 and *G9a-/-* ES cells (Figure 4.1C). As expected, H2K9me2 was reduced by 75% in shG9a-2 single clonal ES cells when compared with shCont ES cells. However, there was no significant reduction of H3K9me1 in shG9a-2 single clonal ES cells and the levels of H3K9me1 in shCont, shG9a-2 and *G9a-/-* ES cells were very similar (Figure 4.1C). This result conflicts with previous observations (Peters et al., 2003; Rice et al., 2003; Tachibana et al., 2002) and suggests that in addition to G9a and GLP, some other histone modifiers are also involved in the establishment and maintenance of H3K9me1. Furthermore, compared with shCont ES cells, the level of H3K9me3 was reduced by 20% in shG9a-2 single clonal ES cells, yet higher than the level detected in *G9a-/-* ES

cells (Figure 4.1C). This result suggests that histone methyltransferases G9a and GLP are partially required for H3K9me3 in ES cells through providing H3K9me2 substrate for the other histone modifiers of H3K9, such as SETDB1 and SUV39H1/2.

To further investigate the levels of H3K9me at ICRs in shG9a ES cells, I carried out ChIP of H3K9me1, H3K9me2 and H3K9me3 at some of the ICRs, *Igf2r*, *Snrpn*, *H19-Igf2*, *Zac1*, *Peg3*, *Peg10*, *Peg13*, *Rasgrf1*, *Kcnq1ot1* and *GnasXL*, in shCont, shG9a-2 and *G9a-/-* ES cells. For H3K9me2 ChIP, *Ndufa* and *Wfdc15* served as negative and positive control regions, respectively. Consistent with Western blotting results, ChIP result of H3K9me2 indicated that compared with shCont ES cells, the level of H3K9me2 reduced more than 70% all the investigated ICRs in shG9a-2 ES cells. In addition, at some of the ICRs, including *Igf2r*, *Snrpn*, *H19-Igf2*, *Zac1* and *Peg3,* the remaining levels of H3K9me2 in shG9a-2 ES cells are even comparable with the levels in *G9a-/-* ES cells (Figure 4.2). These results were consistent with the previous observations (Tachibana et al., 2002; Tachibana et al., 2005) and suggested that knockdown of *G9a* not only reduces the global level of H3K9me2, but also reduces the levels of H3K9me2 at ICRs in ES cells.

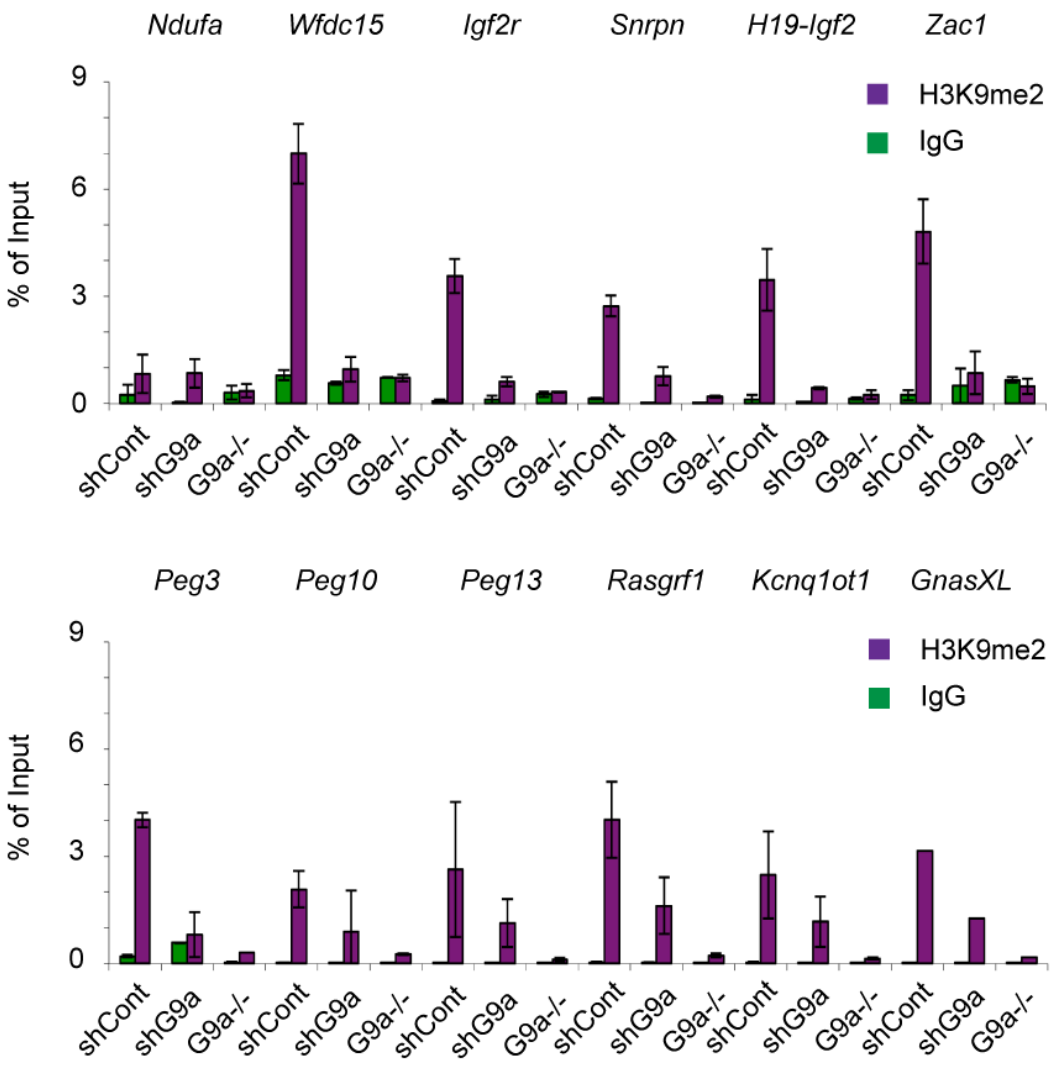


Figure 4.2 H3K9me2 ChIP at imprinted loci in shCont, shG9a and *G9a-/-* ES cells.

Ndufa and *Wfdc15* serve as the negative and positive control regions for H3K9me2, respectively. *H19-Igf2* and *Rasgrf1* are two paternal ICRs, whereas *Igf2r*, *Snrpn*, *Zac1*, *Peg3*, *Peg10*, *Peg13*, *Kcnq1ot1* and *GnasXL* are eight maternal ICRs. The Y-axis represents percentage of enrichment relative to Input DNA. Mouse IgG (green bars) was used as a negative control for H3K9me2 ChIP. Error bars represent 2 times standard deviations of technical triplicates.

In addition, ChIP result of H3K9me1 indicated that compared with shCont ES cells, the levels of H3K9me1 in shG9a ES cells varied between loci, since the levels of H3K9me1 significantly reduced at some of the investigated ICRs but not at the others (Figure 4.3). These results indicated that histone methyltransferases G9a and GLP play a selective role in the maintenance of H3K9me1 at some of the ICRs. However at the other ICRs, some other histone modifiers must exist and play a role in the maintenance of H3K9me1.

For the levels of H3K9me3, ChIP results indicated that compared with shCont ES cells, there was an approximately 50% reduction at all the investigated ICRs in shG9a-2 ES cells, yet the remaining levels of H3K9m3 in shG9a-2 ES cells were still higher than the levels in *G9a-/-* ES cells (Figure 4.4). These results were consistent with previous Western blotting results and together suggested that histone methyltransferases G9a and GLP also contributes to the level of H3K9me3 in ES cells.

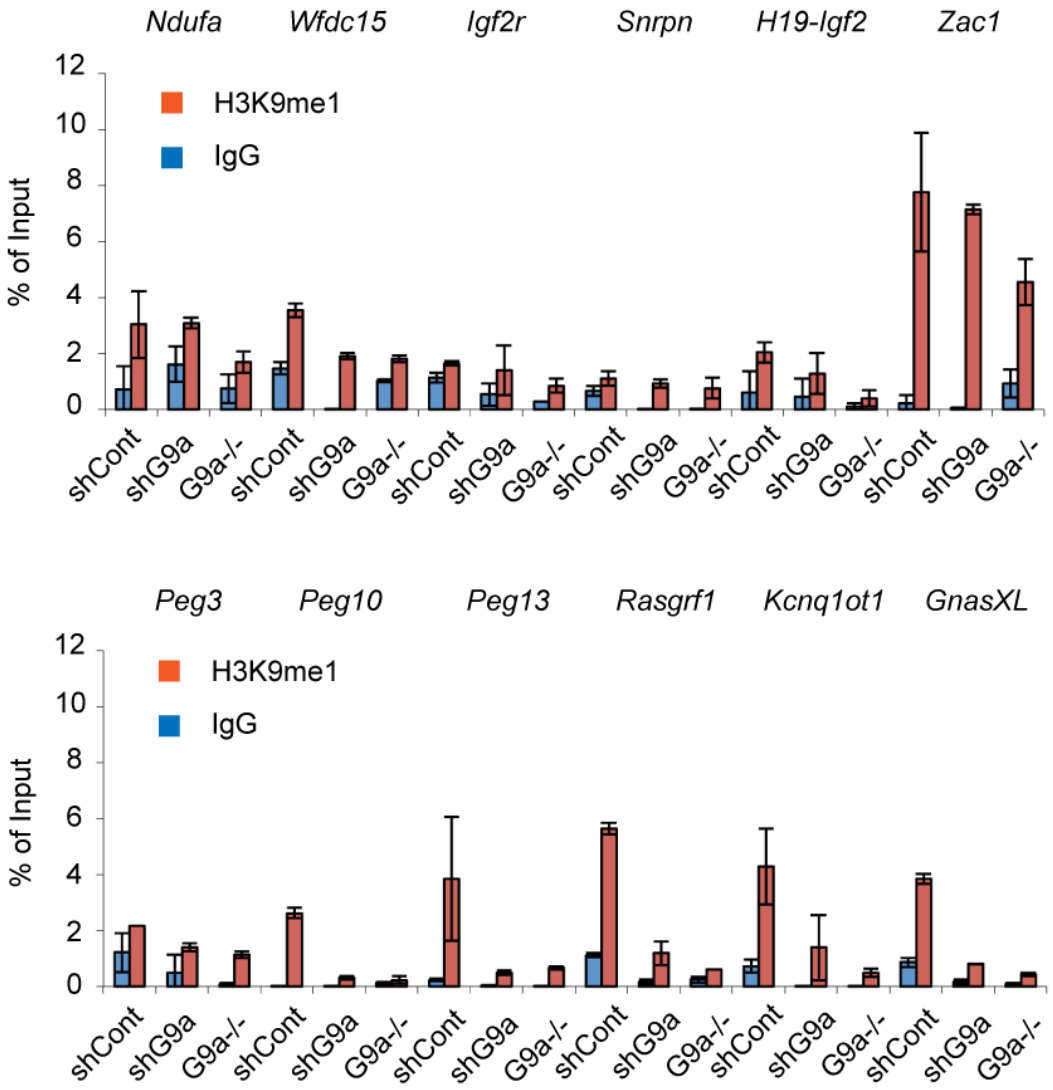


Figure 4.3 H3K9me1 ChIP at imprinted loci in shCont, shG9a and *G9a-/-* ES cells.

Ndufa and *Wfdc15* serve as the negative and positive control regions for H3K9me2, respectively. *H19-Igf2* and *Rasgrf1* are two paternal ICRs, whereas *Igf2r*, *Snrpn*, *Zac1*, *Peg3*, *Peg10*, *Peg13*, *Kcnq1ot1* and *GnasXL* are eight maternal ICRs. The Y-axis represents percentage of enrichment relative to Input DNA. Mouse IgG (blue bars) was used as a negative control for H3K9me1 ChIP. Error bars represent 2 times standard deviations of technical triplicates.

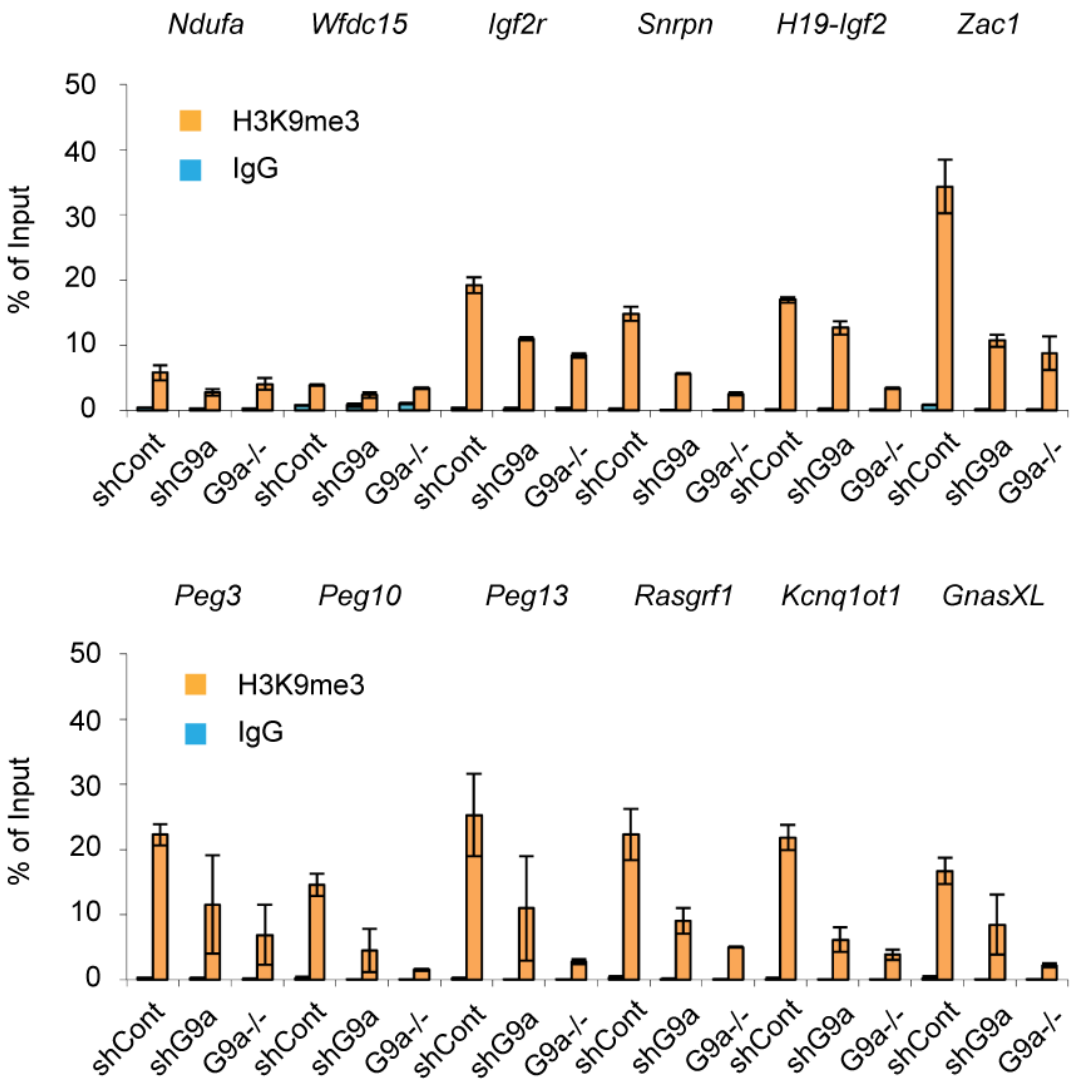


Figure 4.4 H3K9me3 ChIP at imprinted loci in shCont, shG9a and *G9a-/-* ES cells.

Ndufa and *Wfdc15* serve as the negative and positive control regions for H3K9me2, respectively. *H19-Igf2* and *Rasgrf1* are two paternal ICRs, whereas *Igf2r*, *Snrpn*, *Zac1*, *Peg3*, *Peg10*, *Peg13*, *Kcnq1ot1* and *GnasXL* are eight maternal ICRs. The Y-axis represents percentage of enrichment relative to Input DNA. Mouse IgG (blue bars) was used as a negative control for H3K9me3 ChIP. Error bars represent 2 times standard deviations of technical triplicates.

4.2.2 Imprinted DNA methylation is reduced in the shG9a ES cells

In order to investigate the levels of imprinted DNA methylation in shG9a ES cells, I performed bisulfite DNA sequencing at *Igf2r*, *Snrpn* and *H19-Igf2* ICRs in shCont and shG9a ES cells. The results showed that both *Igf2r* and *H19-Igf2* ICRs displayed significant loss of imprinted DNA methylation in shG9a ES cells when compared with shCont ES cells (Figure 4.5). In contrast, there was only 15.6% of loss of imprinted DNA methylation at *Snrpn* ICR (46.5% in shCont and 30.9% in shG9a), which suggests that imprinted DNA methylation is less stable at *Igf2r* and *H19-Igf2* ICRs when the level of G9a is reduced by shRNA.

In addition, to validate bisulfite DNA sequencing results and investigate imprinted DNA methylation further, I carried out 5mC MeDIP at ten ICRs in shCont, shG9a and *G9a-/-* ES cells. Again, *Ndufa* and *Ankrd50* served as negative and positive control regions, respectively. The results showed most of the investigated ICRs, including *Igf2r*, *Snrpn*, *H19-Igf2*, *Zac1*, *Peg3*, *Peg13*, *Kcnq1ot1* and *GnasXL*, imprinted DNA methylation was reduced by approximately 30% in shG9a ES cells when compared to shCont ES cells (Figure 4.6). No significant changes at *Peg10* and *Rasgrf1* loci between shCont and shG9a ES cells might due to relative more CpG sites within these DMRs. Therefore, loss of DNA methylation occurred at a small proportion within these CpG sites could not be discriminated by 5meC MeDIP. However, unlike in *G9a-/-* ES cells, loss of imprinted DNA methylation was only partial in shG9a ES cells. These data are consistent with my bisulfite DNA sequencing experiments at *Igf2r* and *Snrpn* loci (Figure 4.5).

However, compared with shCont ES cells, 5meC MeDIP result showed approximately 30% reduction of DNA methylation at *H19-Igf2* locus in shG9a ES cells

(Figure 4.5), whereas bisulfite DNA sequencing result showed nearly all the imprinted DNA methylation was lost at *H19-Igf2* locus in shG9a ES cells (Figure 4.6). This conflict might due to technical difference between MeDIP and bisulfite DNA sequencing, since methylated CpG sites within the flanking regions of a specific locus might influence the MeDIP, but not the bisulfite DNA sequencing result. Taken together, these data indicate that G9a is essential for the maintenance of imprinted DNA methylation in ES cells.

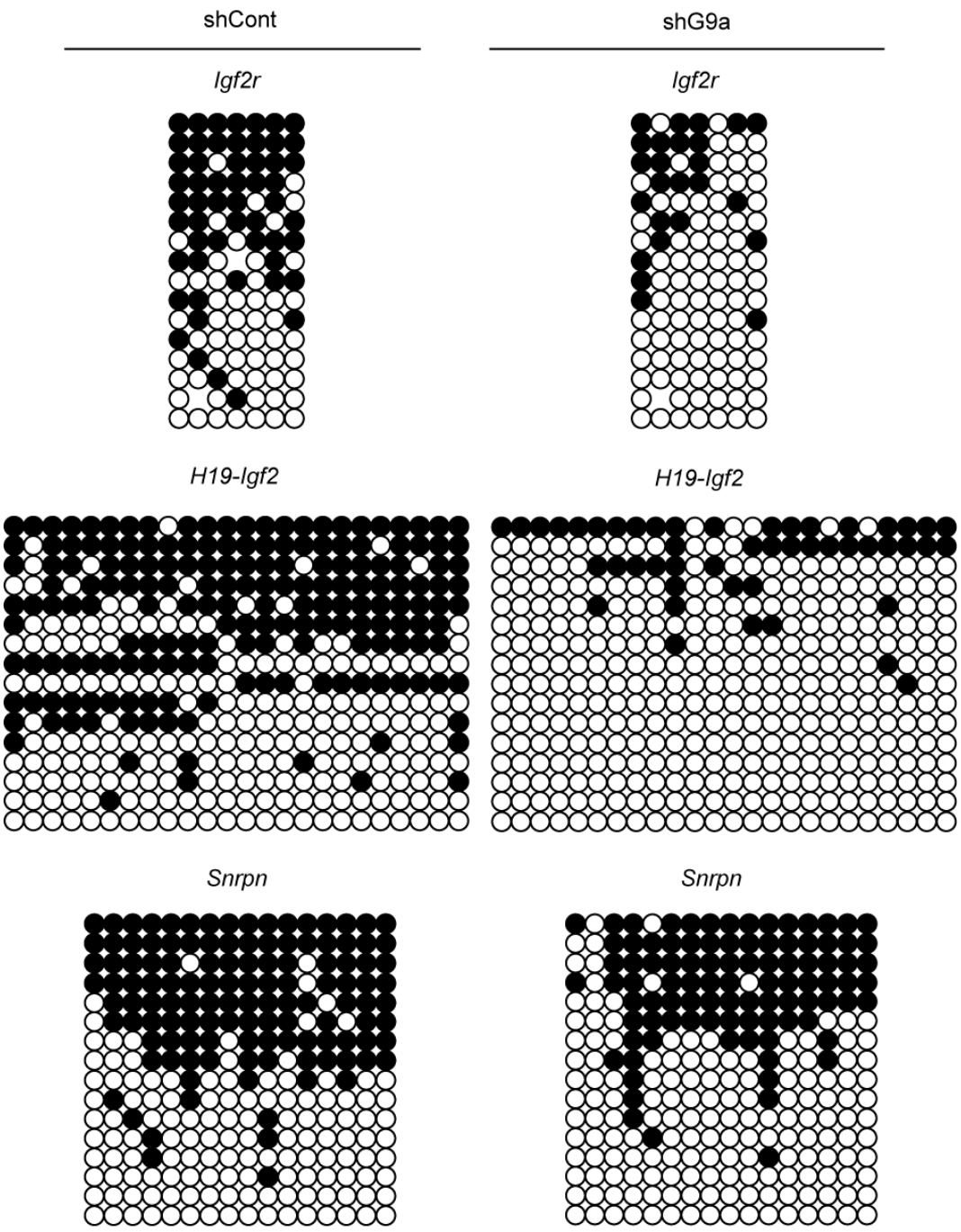


Figure 4.5 DNA methylation is reduced at ICRs in TT2 shG9a ES cells.

Bisulfite DNA sequencing at *Igf2r, Snrpn* and *H19-Igf2* ICRs in ES cells expressing stable either control non-targeting shRNA (shCont) or shRNA that targets G9a (shG9a). The black circles indicate the methylated CpGs, whereas the white circles represent the unmethylated CpGs. Each row represents a single DNA strand.

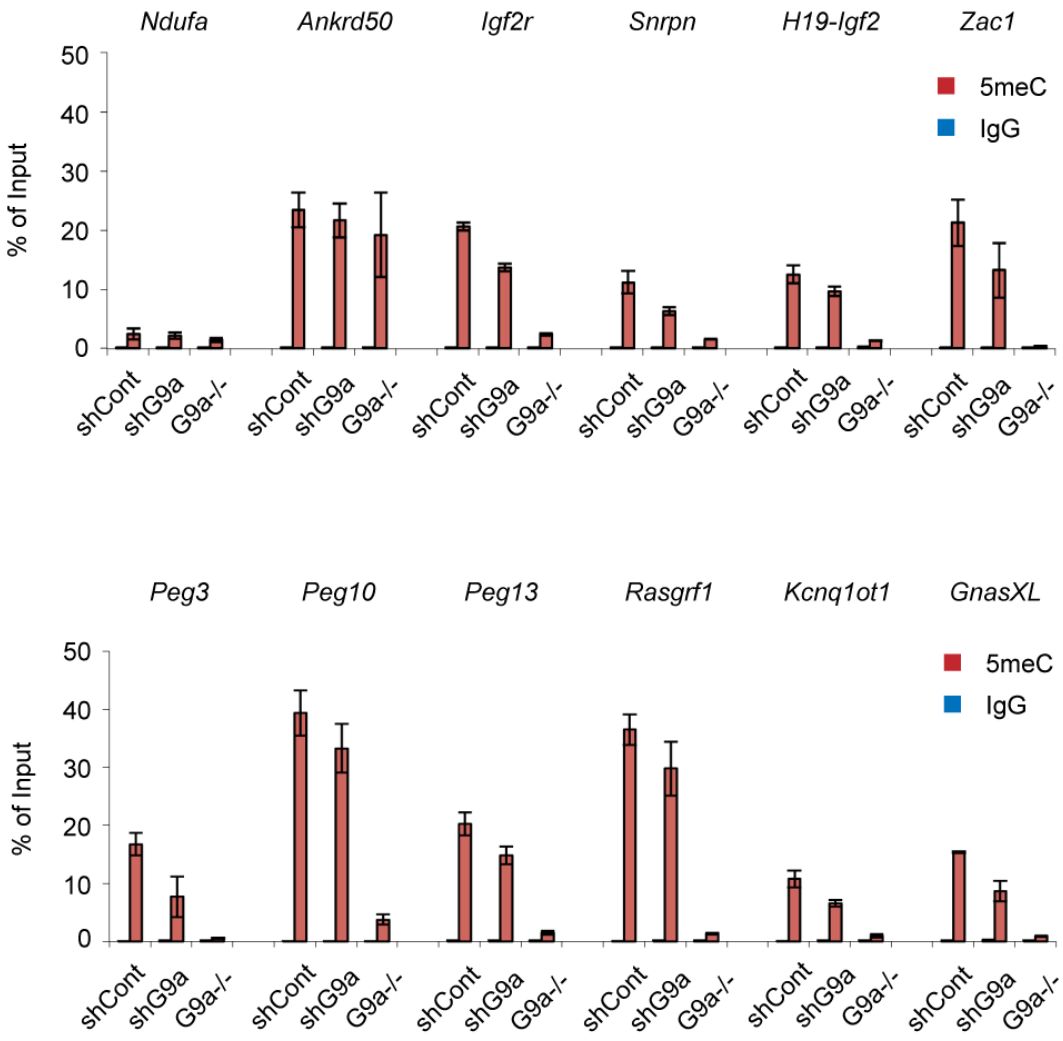


Figure 4.6 Knockdown of G9a reduces imprinted DNA methylation levels in ES cells.

Ndufa and *Ankrd50* serve as the negative and positive control regions, respectively. *H19-Igf2* and *Rasgrf1* are two paternal ICRs, whereas *Igf2r*, *Snrpn*, *Zac1*, *Peg3*, *Peg10*, *Peg13*, *Kcnq1ot1* and *GnasXL* are the maternal ICRs. Y-axis value represents percentage of enrichment normalized to input DNA. IgG (blue bars) was employed as the negative control for 5meC MeDIP. Error bars represent 2 times standard deviations of technical triplicates.

4.2.3 Effect of G9a on imprinted DNA methylation is independent of genetic background of ES cells

To investigate whether the effect of knockdown of *G9a* on imprinted DNA methylation could be reproduced in ES cells with different genetic background, I also stably knocked down *G9a* by shRNA in E14 wild-type ES cells. As previously, I first generated a pool clonal *G9a* knockdown ES cell line in E14 ES cells. Then, I picked six independent shG9a single clonal ES cell lines from pool clonal ES cell line, in which I tested the levels of G9a, GLP and H3K9me2 by Western blotting assays. The results showed that the levels of G9a, GLP and H3K9me2 were reduced in shG9a pool clonal ES cells when compared with shCont ES cells (Figure 4.7A). However, among individual shG9a single clonal cell lines, the levels of G9a, GLP and H3K9me2 varied (Figure 4.7A). I selected shG9a-9 single clonal ES cell line for downstream experiments, as it displayed the most efficient knockdown of *G9a* and the lowest levels of GLP and H3K9me2 (Figure 4.7B).

To investigate the patterns of imprinted DNA methylation in shG9a and shCont E14 ES cells, I carried out bisulfite DNA sequencing at *Ankrd50* promoter and *Igf2r* DMR2 loci (Figure 4.8A). At *Ankrd50* promoter, both shCont and shG9a ES cells displayed comparably high levels of DNA methylation. However, imprinted DNA methylation at *Igf2r* DMR2 was only reduced in shG9a E14 ES cells, but not in shCont ES cells, which indicates that G9a is also required for the maintenance of imprinted DNA methylation in E14 ES cells.

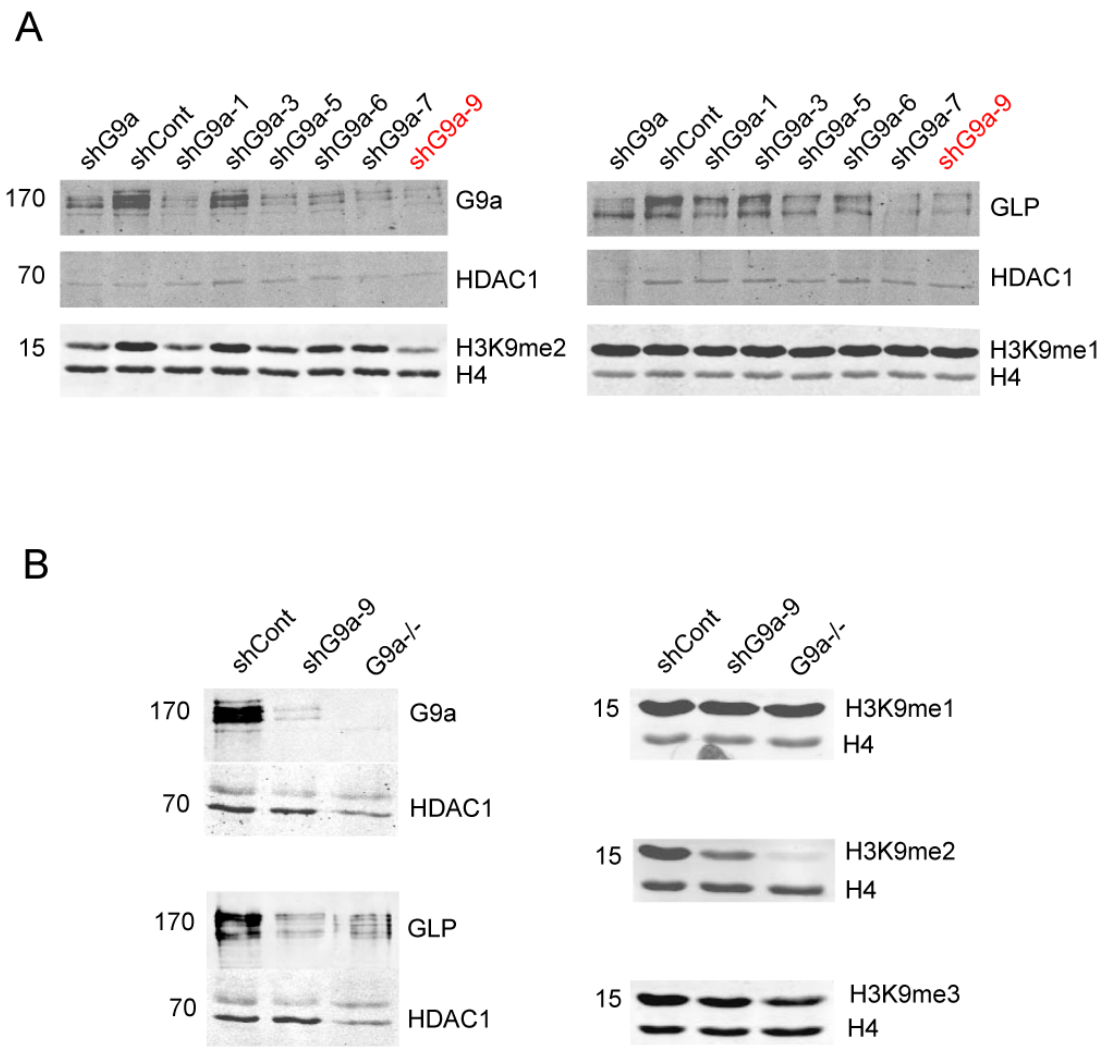


Figure 4.7 Generation of stable shG9a clonal cell line in E14 ES cells.

(**A**) Western blots detecting G9a, GLP and H3K9 methylation levels in shG9a pool clonal cell line, shCont single clonal cell line and six individual shG9a single clonal cell lines. The single clonal cell lines marked in red were used for the further analyses. (**B**) Western blots detecting of G9a and GLP protein and H3K9 methylation levels in shCont single clonal cell line, shG9a single clonal cell line and *G9a-/-* cells. HDAC1 and H4 are served as the loading controls for G9a or GLP and H3K9me, respectively.

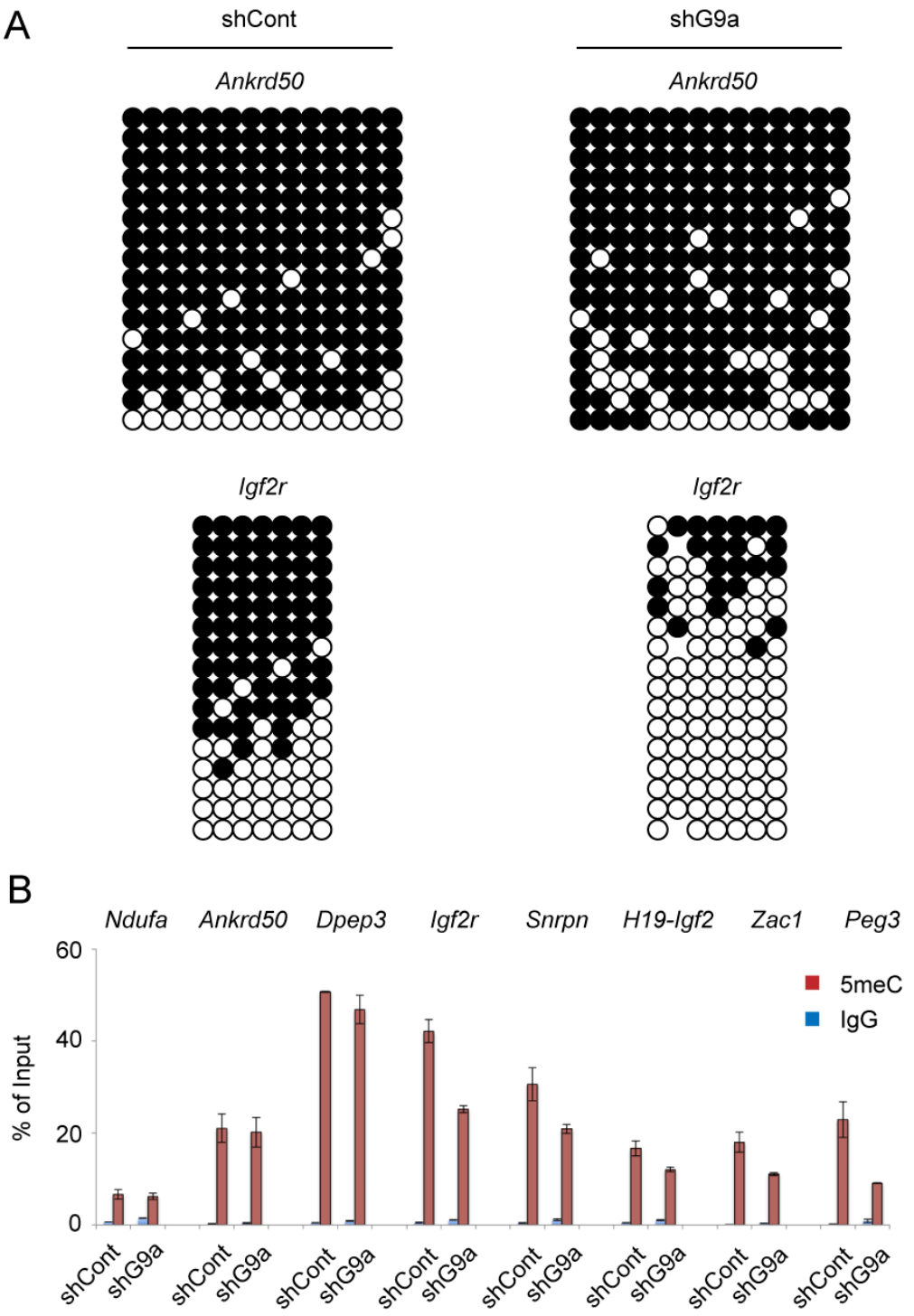


Figure 4.8 Knockdown of G9a in E14 ES cells reduces imprinted DNA methylation.

(**A**) Bisulfite DNA sequencing at *Ankrd50* promoter and *Igf2r* ICR in shCont and shG9a ES cells. The black circles indicate the methylated CpGs, whereas the white circles represent the unmethylated CpGs. Each row represents a single DNA strand. (**B**) 5meC MeDIP at imprinted loci in shCont and shG9a ES cells. *Ndufa* and *Ankrd50* serve as the negative and positive control regions, respectively. The Y-axis represents percentage enrichment relative to input DNA. Mouse IgG (blue bars) is as the negative control for 5meC MeDIP. Error bars represent 2 times standard deviations of technical triplicates.

To validate bisulfite DNA sequencing data and investigate the patterns of imprinted DNA methylation at a larger number of ICRs, I performed 5meC MeDIP in shG9a and shCont E14 ES cells (Figure 4.8B). Again, *Ndufa* and *Ankrd50* served as negative and positive control regions, respectively. As expected, 5meC MeDIP results showed that both shCont and shG9a ES cells displayed hypo- and hyper-methylation at *Ndufa* and *Ankrd50* loci. *Dpep3* was employed as another positive control region, which also showed no changes of DNA methylation according to the promoter microarray data and 5mC MeDIP results in both shCont and shG9a E14 ES cells (Figure 4.8B).

In comparison with shCont ES cells, imprinted DNA methylation at all the investigated ICRs, including *Igf2r*, *Snrpn*, *H19-Igf2*, *Zac1* and *Peg3*, was reduced by about 20-60% in shG9a ES cells (Figure 4.8B). These data are in agreement with bisulfite DNA sequencing results at *Igf2r* DMR2 locus (Figure 4.8A). Whereas some of the DMRs that have been checked in shG9a TT2 ES cells (Figure 4.6), including *Peg10*, *Peg13*, *Kcnq1ot1*, *GnasXL* and *Rasgrf1*, are not investigated here. Taken together, these data confirmed that G9a is essential for the maintenance of imprinted DNA methylation in ES cells. In summary, experiments performed within two independent ES cells lines that have different genetic background consistently demonstrate that reducing levels of G9a by shRNA leads to reduction of imprinted DNA methylation at ICR in ES cells.

4.3 Discussion

Both bisulfite DNA sequencing and 5meC MeDIP experiments clearly demonstrate that knockdown of *G9a* leads to the reduction of DNA methylation at most of the ICRs in ES cells (Figures 4.5, 4.6 and 4.8). Since *G9a-/-* ES cells are derived from wild-type ES cells by gene conversions, which should carry normal patterns of imprinted DNA methylation originally, impairment of imprinted DNA methylation in shG9a and

G9a-/- ES cells confirmed that G9a plays an important role in the maintenance of imprinted DNA methylation.

In addition, previous studies have suggested that even in wild-type ES cells, culturing for a long time *in vitro* could affect the normal patterns of imprinted DNA methylation (Dean et al., 1998; Zvetkova et al., 2005). Since shG9a single clonal ES cells were generated together with shCont single clonal ES cells and all these cells were cultured for approximate the same length of time *in vitro*, reduction of imprinted DNA methylation in shG9a ES cells is not due to extended passaging (Figures 4.5, 4.6 and 4.8). In addition, my data also clearly show that the effect of knockdown of *G9a* on the imprinted DNA methylation is independent of genetic background of ES cells, as both shG9a TT2 and E14 ES cells displayed similar reduction of imprinted DNA methylation when investigated by bisulfite DNA sequencing and 5mC MeDIP.

Since the process of generation of shG9a single clonal ES cell line required approximate 4 weeks, it is unclear that how imprinted DNA methylation is lost form ICRs in shG9a ES cells, by active or passive DNA demethylation? To address this question, *G9a* conditional deficient ES cell line is required and the patterns of imprinted DNA methylation need to be investigated soon after G9a is removed. Alternatively, since active DNA demethylation requires oxidase and/or glycosylase, such as ten-eleven translocation (TET) oxidases and thymine DNA glycosylase (TDG) (Ito et al., 2010), knockdown of *G9a* in the active DNA demethylation-associated enzymes deficient ES cells following by investigation of the patterns of imprinted DNA methylation may shed light on whether loss of the imprinted DNA methylation in *G9a* deficient ES cells through active DNA demethyaltion.

As mentioned above, G9a and GLP are the predominant histone methyltransferases for the establishment and maintenance of H3K9me2 in mammals (Rice et al., 2003; Tachibana et al., 2002). Therefore, it is not surprising that knockdown of *G9a* causes reduction of global H3K9me2 in ES cells. Since my ChIP experiments demonstrate that the levels of H3K9me2 are also reduced at ICRs in shG9a ES cells, it is possible that G9a is somehow directly function at ICRs and deposits H3K9me2, which protects ICRs from either active or passive DNA demethylation in ES cells. This is consistent with study of STELLA (also known as PGC7), a maternal factor, which could recognise and bind to H3K9me2 and protect maternal and paternal ICRs, including *Mest, Peg3, Peg5, Dlk1-Gtl2, H19-Igf2 and Rasgrf1,* from loss of DNA methylation in zygote (Nakamura et al., 2007; Nakamura et al., 2012).

In addition, although previous studies have shown that G9a and its partner GLP also play a role in the establishment and maintenance of H3K9me1 *in vivo* (Tachibana et al., 2002; Tachibana et al., 2005; Ueda et al., 2006), my results suggest that global level of H3K9me1 is not significantly reduced and the levels of H3K9me1 only reduced at some of the ICRs in shG9a ES cells. These observations conflict with previous studies and indicate that histone methyltransferases G9a and GLP are not the only histone modifiers for the establishment and maintenance of H3K9me1 in ES cells. Actually, a recent study has shown that histone methyltransferases PRDM3 and PRDM16 also serve as the histone modifiers for the establishment and maintenance of H3K9me1 in ES cells (Pinheiro et al., 2012). In addition, since non-nucleosomal H3 that contains H3K9me1 also exist in cytoplasma (Loyola et al., 2006), my shG9a ES cells might also undergo loss of H3K9me1 globally. However, by some unknown mechanism, nucleosomal histone H3

within the majority of genomic loci that lost H3K9me1 are replaced by the free H3K9me1 modified histone H3.

Furthermore, my experiments also detected the reduction of H3K9me3 at ICRs in shG9a ES cells. This result is consistent with the previous observation and together suggests that apart from H3K9me2, histone methyltransferase G9a is also required for the maintenance of H3K9me3 *in vivo* (Yokochi et al., 2009). A possible explanation is that H3K9me2 serves as a substrate for the histone methyltransferases of H3K9me3, such as SETDB1 and SUV39H1/2. Therefore reduction of H3K9me3 could be caused by the reduced amount of H3K9me2 in shG9a ES cells. Interestingly, it was recently shown that all the histone modifiers of H3K9me, including G9a, GLP, SETDB1, SUV39H1 and SUV39H2, could form a mega-complex *in vivo* and together regulate the levels of H3K9me in ES cells (Fritsch et al., 2010).

Additionally, my data is also consistent with the previous observations, which indicated that apart from G9a and GLP, other histone methyltransferases also function at the methylated allele at ICRs. It was demonstrated recently that a zinc finger protein ZFP57 is able to bind to the methylated DNA at ICRs and maintains imprinted DNA methylation by recruitment of H3K9 tri-methyltransferase SETDB1, its cofactor KAP1 and DNMTs (Quenneville et al., 2011). Therefore, reduction of H3K9me3 in shG9a ES cells indicates that G9a either upstream of SETDB1 or has an independent role in the maintenance of imprinted DNA methylation in ES cells.

However, since previous studies have shown that H3K9me2 *per se* is sufficient for the regulation of imprinted genes in *Dnmt1-/-* ES cells and G9a protein itself can directly interacts with DNMTs *in vivo* (Esteve et al., 2006; Epsztejn-Litman et al., 2008), reduction of imprinted DNA methylation in shG9a ES cells in my study raised the

questions that how does G9a maintain imprinted DNA methylation in ES cells and is H3K9me2 essential for the maintenance of imprinted DNA methylation by G9a?

Chapter 5 G9a protects the ICRs from loss of DNA methylation independently of its catalytic activity

5.1 Introduction

In mammals, the main modifiers for the establishment and maintenance of H3K9me2 are histone methyltransferases G9a and its partner protein GLP (Tachibana et al., 2005; Ueda et al., 2006). However, apart from the histones, emerging evidence suggests that G9a and GLP could also target to the non-histone proteins, such as p53 and WIZ. For instances, both G9a and GLP have been found to be able to di-methylate the lysine 373 within the tumour suppressor transcription factor protein p53, resulting in an inactive form of p53 (Huang et al., 2010).

In addition to the establishment and maintenance of the H3K9me2, G9a also plays a role in the tumorigenesis. It has been found that G9a is overexpressed in different types of cancers, including leukaemia, hepatocellular carcinoma, prostate carcinoma and the lung cancer (Huang et al., 2010; Kondo et al., 2007; Yamashita et al., 2008). Moreover, it has also been found that knockdown of *G9a* prevents the growth of the cancer cells. Since G9a is also associated with the cocaine addiction (Maze et al., 2010), the mental retardation (Schaefer et al., 2009), the maintenance of HIV-1 latency (Imai et al., 2010) and affects the DNA methylation in mouse ES cells (Dong et al., 2008; Link et al., 2009; Tachibana et al., 2008), increasing investigations are focusing on the roles of G9a and its associated H3K9me2.

In mammalian cells, H3K9me2 often co-localizes with the DNA methylation at the heterochromatin regions. To separately decipher the underlying molecular functions of G9a protein and H3K9me2, selective pharmacological inhibitors of G9a and GLP have

been generated, including the small-molecules BIX01294 and UNC 0638 (Chang et al., 2009; Vedadi et al., 2011). Since the structures of these G9a inhibitors mimic the tail of histone H3, it has been shown that once the cells are treated with these inhibitors, the SET domain within either G9a or GLP is occupied and consequently, these enzymes lose their catalytic activity to maintain the H3K9me2. As shown in several cell lines, BIX01294 reduces both the abundance of the global H3K9me2 and the levels of H3K9me2 at the G9a target regions (Kubicek et al., 2007). However, as BIX01294 is toxic to the cells at the concentrations above 4.1 μM (Kubicek et al., 2007), it is not an ideal inhibitor to investigate the functions of G9a and GLP *in vivo*. By contrast, UNC 0638 displays lower cellular toxicity and the global level of H3K9me2 in the UNC 0638 treated cells is equivalent to the level that observed in the short hairpin RNA (shRNA) knockdown of *G9a* or *GLP* cells (Vedadi et al., 2011). In addition, the levels of H3K9me2 at the G9a target loci are also reduced in the UNC 0638 treated cells (Vedadi et al., 2011). Therefore, UNC 0638 is an ideal G9a inhibitor for investigation the role of G9a by depositing of H3K9me2.

In the previous chapter, my observations have demonstrated that both the levels of DNA methylation and H3K9me2 are impaired at ICRs in *G9a* deficient ES cells (Chapter 4). However, it is still unclear that how does G9a protect ICRs from loss of imprinted DNA methylation in ES cells. One hypothesis was that G9a maintains imprinted DNA methylation by directly recruitment of DNMTs to ICRs, since G9a *per se* was shown to be able to interact with DNMTs directly (Esteve et al., 2006, Epsztejn-Litman et al., 2008). Another hypothesis was that G9a maintains imprinted DNA methylation via the establishment of H3K9me2 at ICRs, since it has been found that nuclear protein STELLA is able to bind to H3K9me2 and protect DNA methylation that is present at the maternal

and paternal ICRs in zygote (Nakamura et al., 2007, 2012). Therefore, the aim of this chapter is to investigate whether G9a directly functions at ICRs and is heterochromatin mark H3K9me2 essential for the maintenance of imprinted DNA methylation in ES cells by G9a.

5.2 Results

5.2.1 H3K9me2 is selectively present at the ICRs on the methylated allele

To investigate whether G9a play a directly role at ICRs in ES cells, I performed precipitation with H3K9me2 antibody followed by bisulfite DNA sequencing of precipitated DNA (Figure 5.1A). I then compared the patterns of DNA methylation at ICRs of H3K9me2 antibody precipitated and input DNA. As expected, all the investigated ICRs, including *Igf2r*, *Snrpn* and *H19-Igf2* ICRs, displayed approximately half of the precipitated DNA fragments were methylated (Figure 5.1B). By contrast, bisulfite DNA sequencing results of H3K9me2 antibody precipitated DNA fragments at these ICRs displayed nearly all the precipitated DNA fragments were methylated (Figure 5.1B). These observations clearly indicate that imprinted DNA methylation is present exclusively on the H3K9me2-marked allele at ICRs in ES cells. In other word, H3K9me2 is selectively established on the methylated allele at ICRs in ES cells.

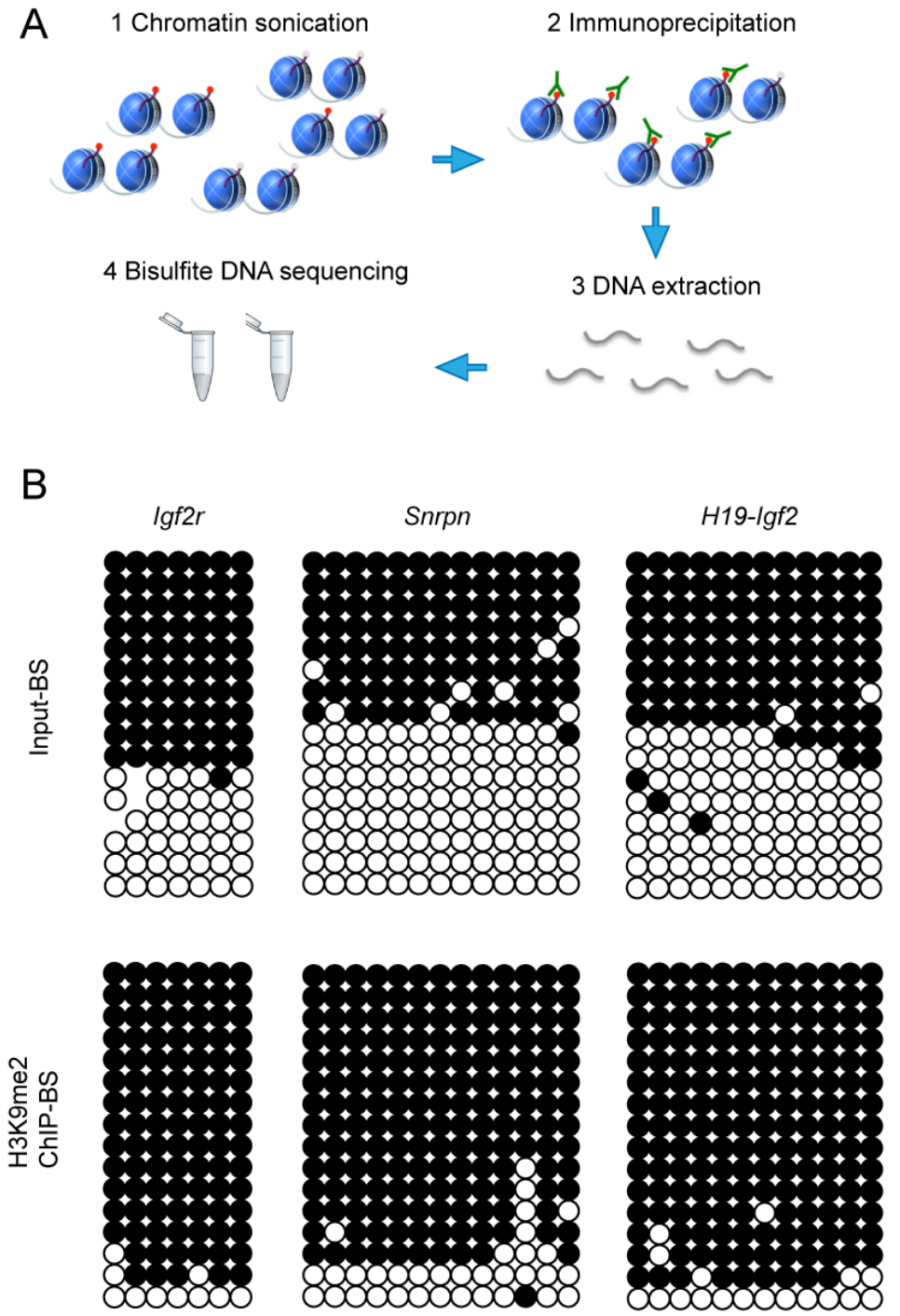


Figure 5.1 H3K9me2 is selectively present on the methylated allele at ICRs.

(A) The schematic view of ChIP for H3K9me2 followed by bisulfite DNA sequencing. (B) Bisulfite DNA sequencing result at *Igf2r*, *Snrpn* and *H19-Igf2* ICRs from H3K9me2 ChIP purified DNA. Input-BS represents bisulfite DNA sequencing carried out using the DNA fragments extracted from sonicated chromatin before H3K9me2 ChIP selection. H3K9me2 ChIP-BS represents bisulfite DNA sequencing carried out using the DNA fragments extracted from sonicated chromatin after H3K9me2 ChIP selection. Each black or white circles indicates a methylated or an unmethylated CpG site. Each row represents a single DNA strand.

Since G9a functions as the main modifier for depositing of H3K9me2 in mammalian cells and either knockout or knockdown of *G9a* leads to the loss of H3K9me2 at ICRs, precipitation with H3K9me2 antibody followed by bisulfite DNA sequencing data suggest that G9a directly functions on the methylated allele at ICRs and G9a seems to protect imprinted DNA methylation by preventing of actions of the DNA demethylase enzymes, such as TET1 and TET2, via establishment of heterochromatin mark H3K9me2. In addition, the presence of H3K9me2 on the methylated allele at ICRs also raises the question that whether H3K9me2 is essential for the maintenance of imprinted DNA methylation by G9a in ES cells.

5.2.2 Treatment of ES cells with G9a/GLP inhibitor leads to the reduction of the global H3K9me2 and the levels of H3K9me2 at the ICRs

To investigate whether H3K9me2 is essential for the maintenance of imprinted DNA methylation by G9a in ES cells, I treated wild-type ES cells with a G9a/GLP inhibitor, UNC0638, and investigated the patterns of imprinted DNA methylation in these cells. The results indicated that, in UNC0638 treated cells, the levels of G9a and GLP remain stable and are comparable with the levels in control ES cells treated with DMSO. However, 2 days afterwards, the level of H3K9me2 is significantly reduced in the inhibitor treated cells (Figure 5.2A). Actually, cells that treated with inhibitor reduced 90% of H3K9me2 within two days and the residual amount of H3K9me2 was comparable to those observed in *G9a-/-* ES cells (Figure 5.2A). In addition, there was no further reduction of H3K9me2 in ES cells that treated with the inhibitor for 4, 6, 8 and 10 days.

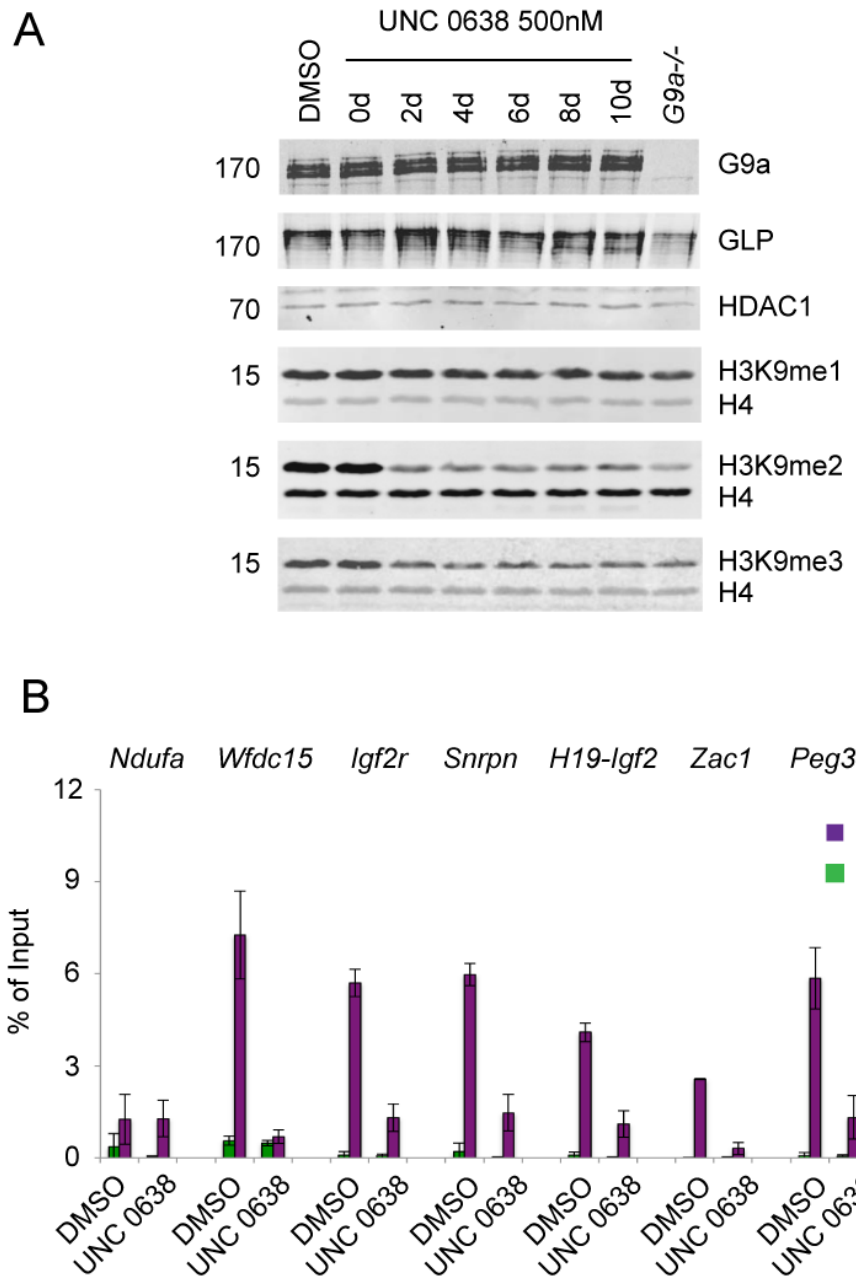


Figure 5.2 Levels of G9a, GLP and H3K9me in UNC 0638 treated ES cells.

(**A**) Western blots detecting G9a, GLP protein and H3K9me levels in ES cells treated with UNC 0638. DMSO and *G9a-/-* represent the wild-type and *G9a-/-* ES cells treated with DMSO for 10 days. 0d, 2d, 4d, 6d, 8d and 10d represent the wild-type ES cells treated with G9a inhibitor (UNC 0638) for 0, 2, 4, 6, 8 and 10 days. HDAC1 and H4 serve as the loading controls for G9a and H3K9me, respectively. (**B**) H3K9me2 ChIP at imprinted loci in UNC 0638 treated ES cells. *Ndufa* and *Wfdc15* serve as the negative and positive control regions, respectively. The Y-axis represents percentage of enrichment relative to Input DNA. Mouse IgG (green bars) was used as a negative control for H3K9me2 ChIP. Error bars represent 2 times standard deviations of technical triplicates.

In contrast to H3K9me2, the global level of H3K9me1 was not significantly reduced in inhibitor treated ES cells when compared with control cells (Figure 5.2A). This result is consistent with my previous observation, which showed that knockdown of *G9a* only reduces the level of H3K9me2, but not H3K9me1, in ES cells. Admittedly, the stability of H3K9me1 indicates that apart from G9a and GLP, some other histone modifiers are involved in the establishment and maintenance of H3K9me1 in ES cells (Pinheiro et al., 2012).

As expected, the level of H3K9me3 was also reduced in UNC 0638 treated ES cells and the level left is comparable with the level observed in *G9a-/-* ES cells (Figure 5.2A). This was expected as the H3K9me2 serves as the substrate for histone methyltransferases of H3K9me3, such as SETDB1 and SUV39H1/2, which could convert H3K9me2 to H3K9me3.

Apart from investigation of the global level of H3K9me2 in UNC 0638 treated ES cells, I also examined the levels of H3K9me2 at ICRs by ChIP. Compared with control ES cells, UNC 0638 treated ES cells displayed reduced levels of H3K9me2 at all the investigated ICRs, including *Igf2r*, *Snrpn*, *H19-Igf2*, *Zac1*, *Peg3* and *Peg10*, as well as at the non-imprinted promoter of G9a target gene *Wfdc15*. In addition, the levels of H3K9me2 at ICRs in UNC 0638 treated ES cells were very similar to those in shG9a ES cells (Figure 5.2B). Taken together, without altering the protein level of G9a, UNC 0638 treated ES cell is a decent model for the downstream investigations about whether G9a protects imprinted DNA methylation via its catalytic activity.

5.2.3 Maintenance of the imprinted DNA methylation does not require the catalytic activity of G9a and GLP

To investigate whether H3K9me2 is required for the maintenance of imprinted DNA methylation in ES cells, I investigated the patterns of imprinted DNA methylation at ICRs in both UNC 0638 and DMSO treated control ES cells by 5meC MeDIP and bisulfite DNA sequencing. I first examined the patterns of imprinted DNA methylation at *Igf2r*, *Snrpn*, *H19-Igf2*, *Zac1*, *Peg3* and *Peg10* ICRs by 5meC MeDIP in control and UNC 0638 treated ES cells at day 0, 4, 8 and 12 after addition of inhibitor. For 5meC MeDIP, *Ndufa* and *Dpep3* promoters served as negative and positive control regions, respectively. The MeDIP experiments clearly showed that there is no reduction of imprinted DNA methylation in ES cells that have been treated with UNC 0638 (Figure 5.3). In addition, DNA methylation that present at *Dpep3* promoter also remained stable. I further confirmed 5meC MeDIP experiments by bisuflite DNA sequencing at *Igf2r* and *Snrpn* ICRs in ES cells that treated with UNC 0638 for 12 days (Figure 5.4A).

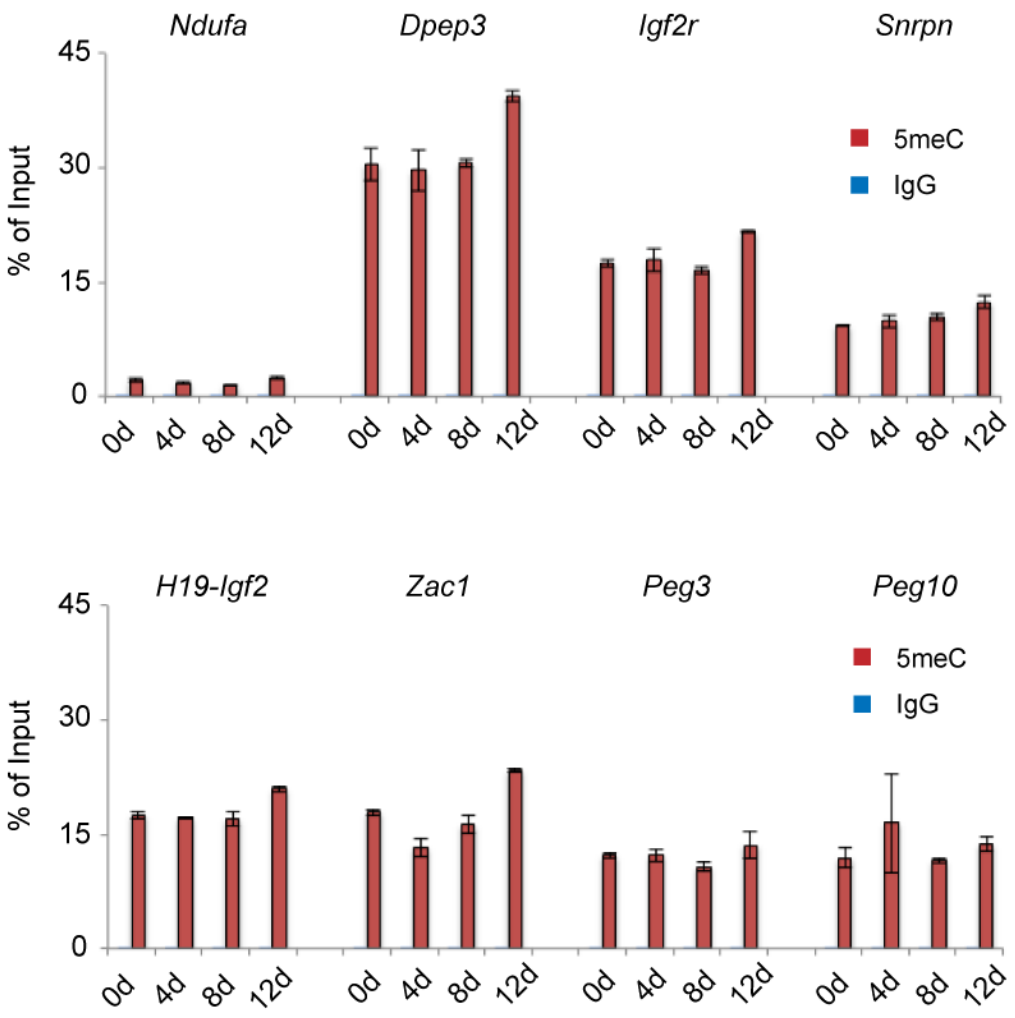


Figure 5.3 DNA methylation remains stable at the ICRs in UNC 0638 treated ES cells.

5meC MeDIP at ICRs in ES cells treated with UNC 0638 for 0, 4, 8 and 12 days. *Ndufa* and *Dpep3* serve as the negative and positive control regions, respectively. The Y-axis value represents percentage of enrichment relative to input DNA. Mouse IgG (blue bars) was used as a negative control for 5meC MeDIP. Error bars represent 2 times standard deviations of technical triplicates.

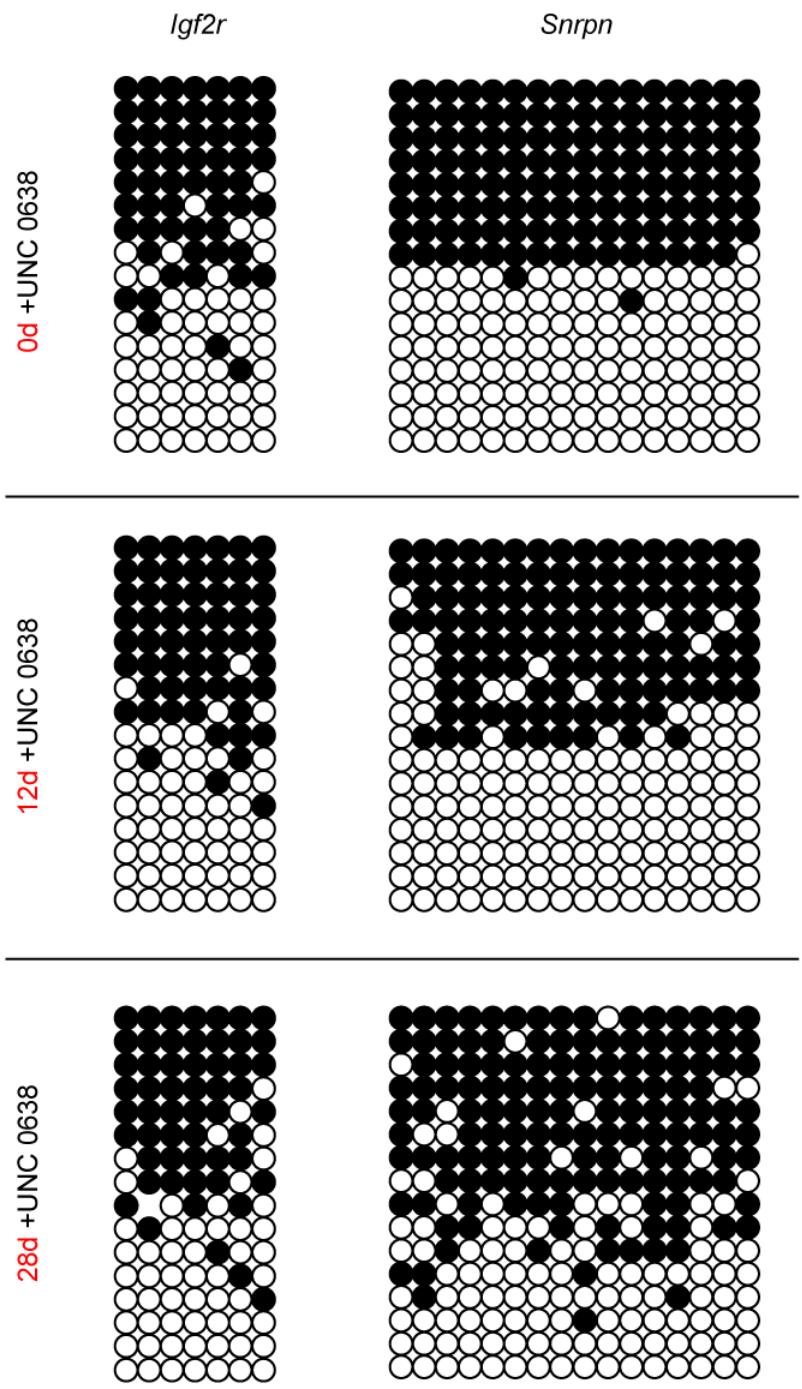


Figure 5.4 Stable DNA methylation patterns at the ICRs in UNC 0638 treated ES cells.

Igf2r and *Snrpn* are two maternally methylated ICRs, which display no reduction of DNA methylation in UNC 0638 treated ES cells. 0d + UNC 0638 represents WT ES cells before treatment with UNC 0638; 12d + UNC 0638 represents WT ES cells treated with UNC 0638 for 12 days; 28d + UNC 0638 represents WT ES cells treated with UNC 0638 for 28 days. Each black or white circle indicates a methylated or an unmethylated CpG sites, respectively. Each row represents a single DNA strand.

Since generation of single clonal shG9a ES cells took approximately 4 weeks, it was possible that the reduced level of H3K9me2 induced by G9a/GLP inhibitor for 12 days is not sufficient for the loss of imprinted DNA methylation in ES cells. Therefore, I treated ES cells with UNC 0638 for 28 days and performed bisulfite DNA sequencing at *Igf2r* and *Snrpn* ICRs. These results clearly showed that even treated with inhibitor for 28 days, imprinted DNA methylation was still stably maintained in ES cells (Figure 5.4). These data demonstrate that G9a/GLP complex is essential for the maintenance of imprinted DNA methylation in ES cells, while its catalytic activity and heterochromatin mark H3K9me2 are completely dispensable.

5.2.4 G9a and GLP bind to the chromatin independently of H3K9me2

The data described above suggest that G9a is likely to maintain imprinted DNA methylation by directly recruitment of DNMTs to ICRs in ES cells. Since H3K9me2 is dramatically reduced in UNC 0638 treated ES cells and both of ANK domains within G9a/GLP complex are involved in recognising and binding to H3K9me1/2 (Collins et al., 2008), G9a might dissociate or partially dissociate from chromatin in UNC 0638 treated ES cells.

To investigate whether G9a still binds to chromatin in UNC 0638 treated ES cells and further ask whether H3K9me2 is essential for the chromatin loading of G9a *in vivo*, I carried out step-wise salt extraction of nuclear proteins from ES cells with buffers containing increasing concentrations of NaCl - 100, 200, 300, 400 and 500 mM, the last one preceded by benzonase nuclease treatment (Figure 5.5A). I then investigated dissociation patterns of G9a and other control chromatin binding proteins by Western blotting assays.

I found that there were more STELLA dissociated from chromatin at 300mM NaCl concentration when compared to other NaCl concentration fractions in wild-type ES cells that were treated with DMSO, whereas in UNC 0638 treated and *G9a-/-* ES cells, I found that there were more STELLA dissociated from chromatin at 100 mM NaCl concentration when compared to other NaCl concentration fractions (Figure 5.5B). This observation is consistent with previous study, which suggested that the chromatin loading patterns of STELLA depends on the presence of H3K9me2 (Nakamura et al, 2007, Nakamura et al, 2012). As expected, histones dissociated from chromatin at 500mM NaCl in all the investigated ES cell lines (Figure 5.5B). Taken together, these results indicate that step-wise salt extraction of nuclear proteins worked well.

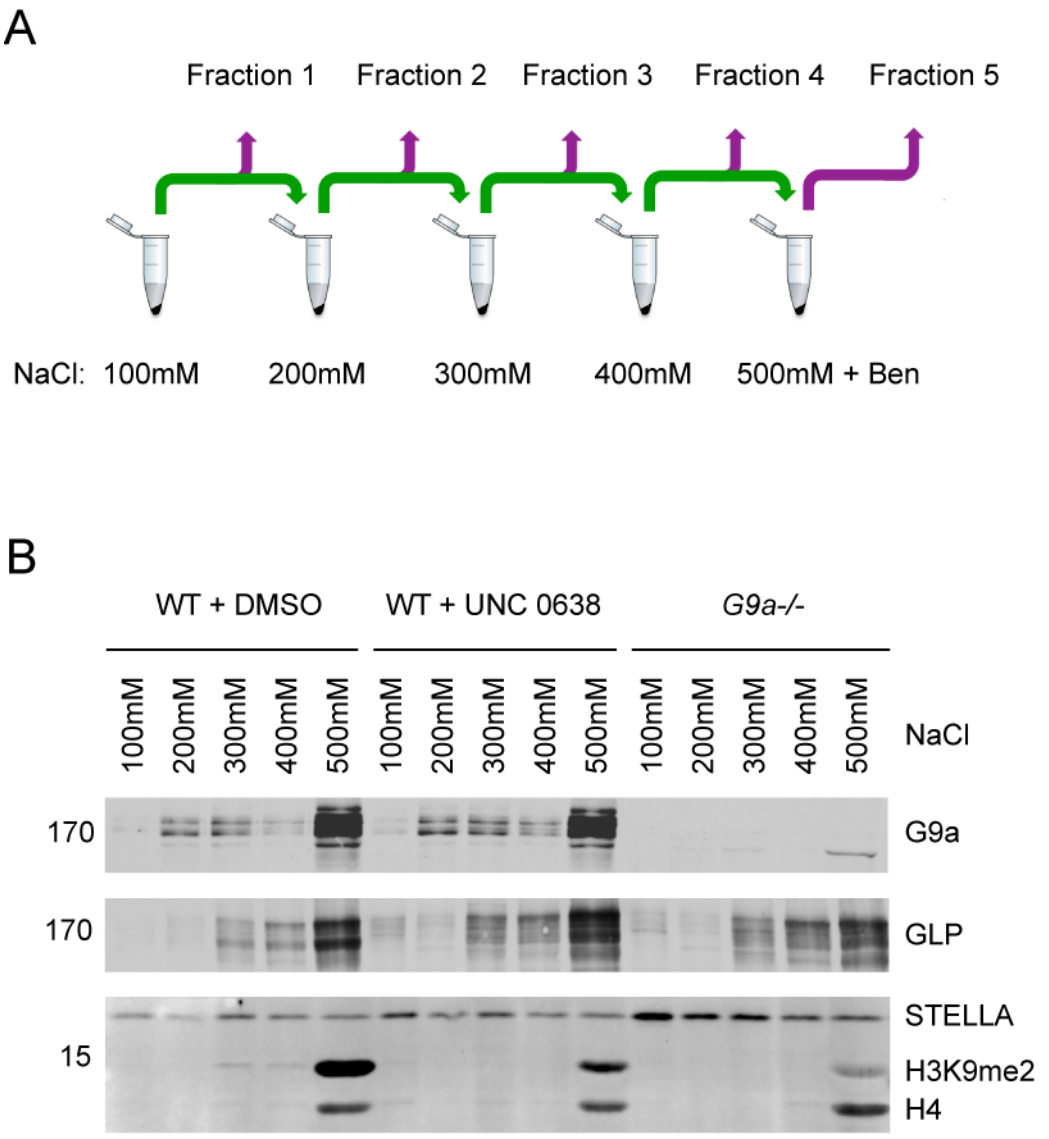


Figure 5.5 The majority of G9a binds to the chromatin independently of H3K9me2.

(**A**) The schematic view of step-wise salt extraction of nuclear proteins. "Ben" represents Benzonase. (**B**) Western blots on protein fractions from step-wise salt extractions probed with antibodies against G9a, GLP, STELLA, H3K9me2 and H4. The nuclear proteins were extracted from wild-type ES cells treated with DMSO for 12 days, wild-type ES cells treated with G9a inhibitor (UNC 0638) for 12 days and the *G9a-/-* ES cells. STELLA is a H3K9me2 binding protein. H4 is a component of nucleosome and serves as a control protein for the step-wise extraction of nuclear proteins.

Interestingly, only a small proportion of G9a and GLP dissociated from chromatin at low salt concentration in UNC 0638 treated ES cells when compared with control ES cells (Figure 5.5B), whereas the vast majority of G9a and GLP dissociated from chromatin at high salt concentration (500 mM NaCl). This observation suggests that histone methyltransferases G9a and GLP are still able to tightly bind to chromatin when the level of H3K9me2 is reduced. In addition, this observation also suggests that G9a and GLP are likely to maintain imprinted DNA methylation by directly recruitment of DNMTs to ICRs in ES cells. However, to investigate whether interaction between G9a and DNMTs is essential for the maintenance of imprinted DNA methylation in ES cells, further experiments have to be carried out.

5.3 Discussion

In this chapter, I first carried out precipitation with H3K9me2 antibody followed by bisulfite DNA sequencing at ICRs and found that imprinted DNA methylation is exclusively present on the H3K9me2-marked allele at ICR in ES cells (Figure 5.1B). Since the majority of global H3K9me2 in mammalian cells is established by histone methyltransferases G9a and GLP (Tachibana et al., 2005; Ueda et al., 2006), this observation indicated that G9a/GLP complex either could discriminate the methylated alleles from the unmethylated ones and directly bind to the methylated allele at ICRs or the unmethylated allele at ICRs are protected from binding of G9a/GLP complex. Since a C2H2 zinc finger protein, WIZ, is also present in G9a/GLP complex and knockdown of *Wiz* leads to instability of the G9a/GLP complex *in vivo* (Tachibana et al., 2005; Ueda et al., 2006), WIZ might function as a potential methylated DNA binding protein and recruit G9a/GLP complex to the methylated allele at ICRs. In addition, long non-coding RNAs (lncRNAs) might also play a role in the recruitment of G9a/GLP complex to the

methylated allele at ICRs, since previous studies have shown that G9a protein *per se* could interact with mono-allelically expressed imprinted lncRNAs, such as *Airn* (Nagano et al., 2008) and *Kcnq1ot1* (Pandey et al., 2008).

Moreover, since G9a/GLP complex is able to recognize and bind to the modified N-terminal tail of histone H3 via ankyrin repeat (ANK) domains, it would be interesting to investigate whether the unmethylated allele at ICRs contains any specific histone modifications, which might prevent the binding of G9a/GLP complex to chromatin. One study has revealed that in patient-derived cells, the asymmetric distribution of histone modifications on the paternally and maternally derived chromosomes is impaired. In normal cells, H3K9me3 and H4K20me3 exist on the methylated allele, whereas H3K4me2, H3K27me3 and H3K9Ac present on the unmethylated allele at *IGF2-H19* ICR. However, in patient derived cells, H3K9me3 and H4K20me3 biallelically exist in the BWS and H3K4me2, H3K27me3 and H3K9Ac biallelically present in the SRS (Nativio et al., 2011).

I next investigated whether H3K9me2 plays a role in the maintenance of imprinted DNA methylation by G9a in ES cells. I employed a G9a/GLP specific pharmacological inhibitor (UNC 0638) and found that imprinted DNA methylation was still stably maintained when the level of H3K9me2 was dramatically reduced (Figure 5.3 and 5.4). This observation suggests that G9a protein, rather than heterochromatin mark H3K9me2, is essential for the maintenance of imprinted DNA methylation in ES cells. Since previous studies have shown that G9a protein itself could interact with DNMTs via its ANK domain (Esteve et al., 2006, Epsztejn-Litman et al., 2008) and precipitation with H3K9me2 antibody followed by bisulfite DNA sequencing results suggested that G9a could directly bind to the methylated allele at ICRs, it is likely that G9a maintains

imprinted DNA methylation by directly recruitment of DNMTs to ICRs in ES cells. To further investigate whether ANK domain within the G9a protein is essential for the maintenance of imprinted DNA methylation, patterns of imprinted DNA methylation need to be investigated in ES cells that only expressing an ANK domain deleted form of G9a.

In addition, since previous studies have shown that G9a protein *per se* could interact with DNMT3A and DNMT3B in ES cells (Epsztejn-Litman et al., 2008), it is necessary to further investigate which DNMT is essential for the maintenance of imprinted DNA methylation by G9a in ES cells. Actually, it has also been found that G9a protein itself could also interact with DNMT1 in cancer cells (Espada et al., 2004). Therefore, it is also possible that G9a could interact with DNMT1 in ES cells and maintains the imprinted DNA methylation through recruitment of DNMT1 to ICRs in ES cells (Espada et al., 2004). To address these questions, co-immunoprecipitation (Co-IP) experiments of G9a and DNMTs and chromatin immunoprecipitation (ChIP) assays of different DNMTs at ICRs are required in both wild-type and *G9a* deficient ES cells.

Furthermore, I preformed step-wise salt extraction of nuclear proteins and combined with Western blotting assays and found that the majority of G9a still tightly binds to chromatin in G9a/GLP inhibitor treated ES cells (Figure 5.5B). This observation indicates that G9a is able to bind to chromatin independently of H3K9me2. In addition, this observation also confirms that chromatin binding of G9a/GLP complex is essential for the maintenance of imprinted DNA methylation in ES cells.

As described in Chapter 3 and 4, I have found that H3K9me1 is not significantly reduced in *G9a* deficient ES cells (Figure 3.4A, 4.1C and 4.7B). In addition, I also showed that H3K9me2 is present at ICRs (Figure 4.2). Considering that G9a/GLP

complex could recognise and bind to H3K9me2 as well as to H3K9me1 via a loop between ankyrin repeat four and five (Collins et al., 2008), it is possible that when the level of H3K9me2 is reduced, G9a/GLP complex binds to chromatin, including ICRs, by recognising of H3K9me1.

Actually, it has been found that four amino acids within ANK domain of G9a, including W839, W844, W877 and E847, are essential for the interaction between G9a or GLP and H3K9me1 or H3K9me2 (Collins et al., 2007). To further investigate whether H3K9me1 is essential for the recruitment of G9a/GLP complex to ICRs when the level of H3K9me2 is reduced, patterns of imprinted DNA methylation need to be investigated in the ES cells that only expressing an ANK domain mutant form of G9a, which is unable to bind to the methylated tail of histone H3. On the other hand, since zinc finger protein WIZ serves as another stable component within G9a/GLP complex (Tachibana et al., 2005; Ueda et al., 2006), it is possible that WIZ plays a role in the recruitment of G9a/GLP complex to ICRs in ES cells. Therefore, it would be necessary to investigate whether WIZ is able to bind to DNA and is there any sequence specific motif within ICRs that WIZ prefers to bind to.

Chapter 6 G9a maintains imprinted DNA methylation via its ANK domain and directly recruits DNMTs to ICRs

6.1 Introduction

G9a and its partner protein GLP serve as the main writers of H3K9me1 and H3K9me2 in mammalian cells and G9a/GLP complex is the predominant form of either G9a or GLP that exist *in vivo* (Tachibana et al., 2005; Ueda et al., 2006). Biochemically, the catalytic SET domain of G9a alone could bind efficiently to un-modified and H3K9me1-marked histone H3 tails (Collins et al., 2008). Since full-length G9a displays higher affinity for H3K9me1 and H3K9me2 than for unmodified histone H3 tail (Collins et al., 2008), it is believed that apart from SET domain, there must be another functional domain within G9a protein, which is essential for recognising and binding to H3K9me1 and H3K9me2.

Actually, previous study has shown that G9a contains another functional domain that located from 730aa to 965aa, named ankyrin repeat (ANK) domain, including six identical helix-turn-helix-β-turn repeats (Collins et al., 2008). In addition, it has also been shown that the fourth and the fifth helix-turn-helix-β-turn repeats within ANK domain is essential for recognising and binding to H3K9-methylated histone H3 tail. By contrast, they display very low affinity to H3K4-methylated, H3K27-methylated or unmethylated histone H3 tails (Collins et al., 2008). In order to further determine which amino acids within the ANK domain are essential for the interaction between G9a and H3K9me2-modified H3 tail, structural analysis has been performed. This study showed that three tryptophans (Try839, Try844 and Try877) and one glutamine acid residue (Glu847) within ANK domain of G9a form a hydrophobic cage, which is able to bind to H3K9me2 peptide. In addition, it has also been found that replacement of any one of these amino

acids that involved in methyl-lysine cage formation could abrogate the interaction between G9a and H3K9me2-modified H3 tail (Collins et al., 2007).

As discussed previously, G9a and GLP are the predominant histone lysine euchromatic methyltransferases of H3K9me2 in mammalian cells (Tachibana et al., 2002; Peters et al., 2003; Rice et al., 2003). Interestingly, apart from the levels of H3K9me2, the levels of DNA methylation is also reduced in either G9a or GLP deficient ES cells and G9a/GLP associated DNA methylation can be partially restored in G9a-null cells by re-expression of a catalytically inactive G9a (Dong et al., 2008; Epsztejn-Litman et al., 2008; Tachibana et al., 2008). In addition, G9a has also been shown to interact with DNMT1 at replication forks in human cells (Esteve et al., 2006). These observations suggest a role for G9a in the maintenance of correct patterns of DNA methylation and H3K9me2 during DNA replication. In addition to DNMT1, G9a also interacts with DNMT3A and DNMT3B via its ANK domain in mouse ES cells (Epsztejn-Litman et al., 2008). Since DNMT3A and DNMT3B also play a role in the maintenance of DNA methylation (Liang et al., 2002), a role for G9a in coupling the maintenance of DNA methylation with heterochromatin has been proposed.

In the previous chapter, I showed that G9a maintains imprinted DNA methylation independently of its catalytic activity and G9a could bind to chromatin when the level of H3K9me2 is reduced. However, since G9a inhibitor-treated ES cells still contain H3K9me1 and both the ANK domains within G9a and GLP could bind to H3K9me1, it is possible that G9a binds to ICRs and maintains imprinted DNA methylation by recognising H3K9me1 in the G9a inhibitor treated ES cells. To investigate whether H3K9me1 is essential for the maintenance of imprinted DNA methylation in G9a inhibitor treated ES cells, I aimed to generate a mutant form of G9a, which is unable to

bind to methylated histone H3 tail and investigate the patterns of imprinted DNA methylation after removal of endogenous G9a. To investigate whether G9a maintains imprinted DNA methylation by direct recruitment of DNMTs to ICRs in ES cells, I also aimed to generate a mutant form of G9a that carrying ANK domain deletion and investigate whether imprinted DNA methylation is affected after endogenous G9a is removed.

6.2 Results

6.2.1 Mutant forms of G9a are stably expressed in ES cells after the endogenous G9a was knocked down

My previous observations have shown that G9a is essential for the maintenance of imprinted DNA methylation in ES cells and once imprinted DNA methylation reduced, neither wild-type nor catalytically inactive G9a can restore imprinted DNA methylation at ICRs in *G9a* deficient ES cells. Therefore, to investigate the functions of mutant forms of G9a (the ANK domain mutation and deletion) on imprinted DNA methylation, it was essential to first express the mutant forms of G9a in wild-type ES cells and then remove of endogenous G9a.

To establish ES cell lines of interest, expression plasmids, which encode for shRNA resistant forms of G9a that carrying either the mutations in H3K9me2 binding amino acids or the ANK domain deletion, first had to be generated. The strategy for generation of these expression plasmids is shown in Figure 6.1. First, I generated an shRNA resistant form of wild-type G9a cDNA by replacement of seven nucleotides corresponding to the shRNA targeted sequences. These mutations were silent and should not change the sequences of amino acids within G9a protein. Next, I mutated and deleted

ANK domain in shRNA resistant form of wild-type G9a cDNA. Finally, I confirmed these mutations/deletions within G9a cDNA by DNA sequencing (Figure 6.2).

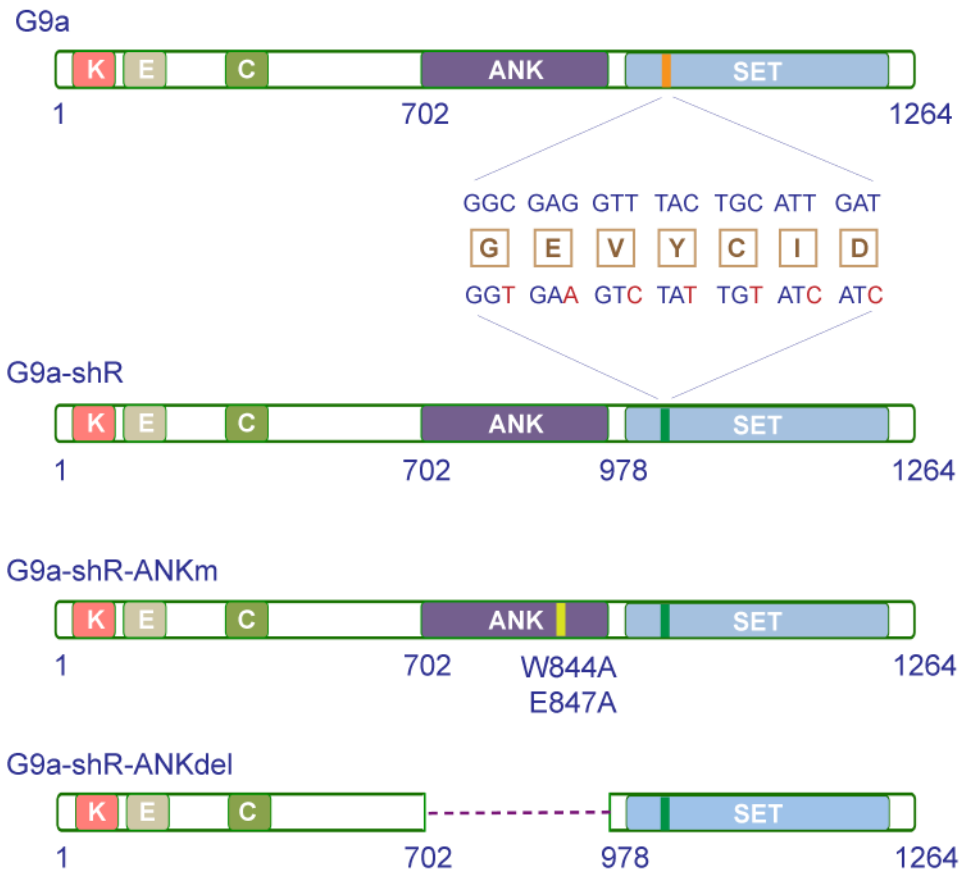


Figure 6.1 The schematic view of mutations and deletions introduced into G9a cDNA and protein.

G9a-shR represents shRNA resistant form of G9a carrying seven base pair substitutions that do not change the amino acid sequence of G9a; G9a-shR-ANKm represents shRNA resistant form of G9a containing Ankyrin repeats domian (ANK) point mutations; G9a-shR-ANKdel represents shRNA resistant form of G9a containing Ankyrin repeats domian (ANK) deletion. K is a lysine rich region; E is a glutamic acid-rich region; C is a cysteine-rich region; ANK represents the Ankyrin repeats domian; SET represents the SET domain. W844A represents a tryptophan to an alanine mutation at 844 residue; E847A represents a glutamic acid to an alanine mutation at 847 residue.

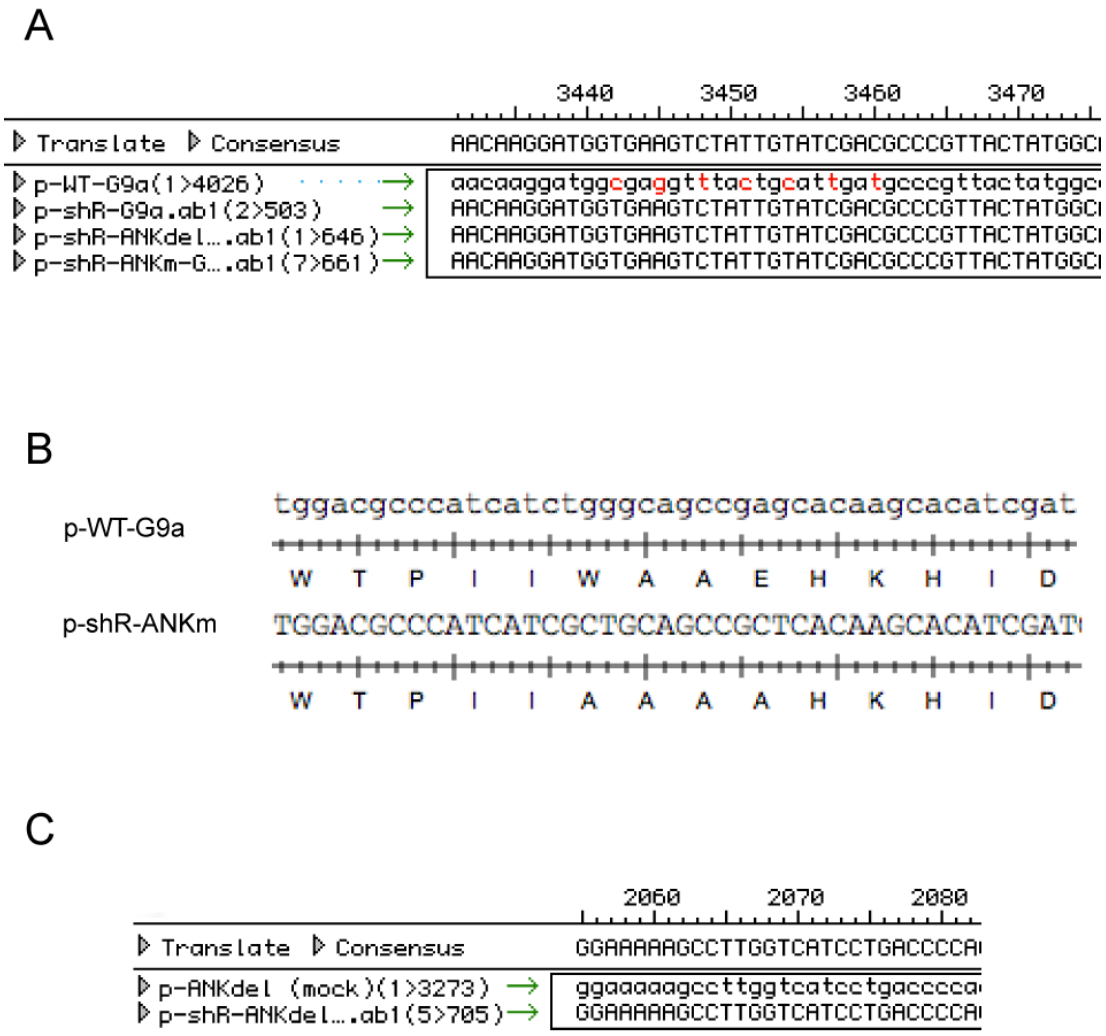


Figure 6.2 Sequencing confirms the ANK domain mutations that introduced into G9a cDNA.

(A) Sequencing confirms the short hairpin RNA (shRNA) resistant mutations that introduced into different ANK domain mutant forms of G9a cDNA. (B) Sequencing confirms the presence of W844A and E847A mutations that encoded by the G9a cDNA. (C) Sequencing confirms the ANK domain deletion that encoded by the G9a cDNA. p-ANKdel (mock) represents the ANK domain deleted form of G9a expressing plasmid, which is virtually generated by the DNASTAR software.

Next, these expression plasmids were linearized and transfected into wild-type ES cells (see material and method) and Western blotting assays were performed in order to investigate the expression of exogenous mutant forms of G9a. These analyses showed that all the exogenous mutant forms of G9a were stably expressed in wild-type ES cells (Figure 6.3A). In addition, GLP and H3K9 methylation levels were also stable in shRNA resistant (shR), ANK domain mutant (ANKm) and ANK domain deleted (ANKdel) ES cell lines (Figure 6.3A).

To specifically investigate the effect of exogenous forms of G9a on imprinted DNA methylation, I then knocked down of endogenous *G9a* in ES cells that expressing shR, ANKm and ANKdel forms of G9a. The mRNA levels of endogenous G9a in established clonal cell lines were tested by quantitative PCR (q-PCR) with primers that anneal across the mutated shRNA targeted sequences, which differs between endogenous and shRNA resistant exogenous forms of G9a. The best three endogenous *G9a* knockdown clonal cell lines for each shR, ANKm and ANKdel expressing ES cells were selected (Figure 6.3B) and investigated for the levels of G9a, GLP and H3K9 methylation.

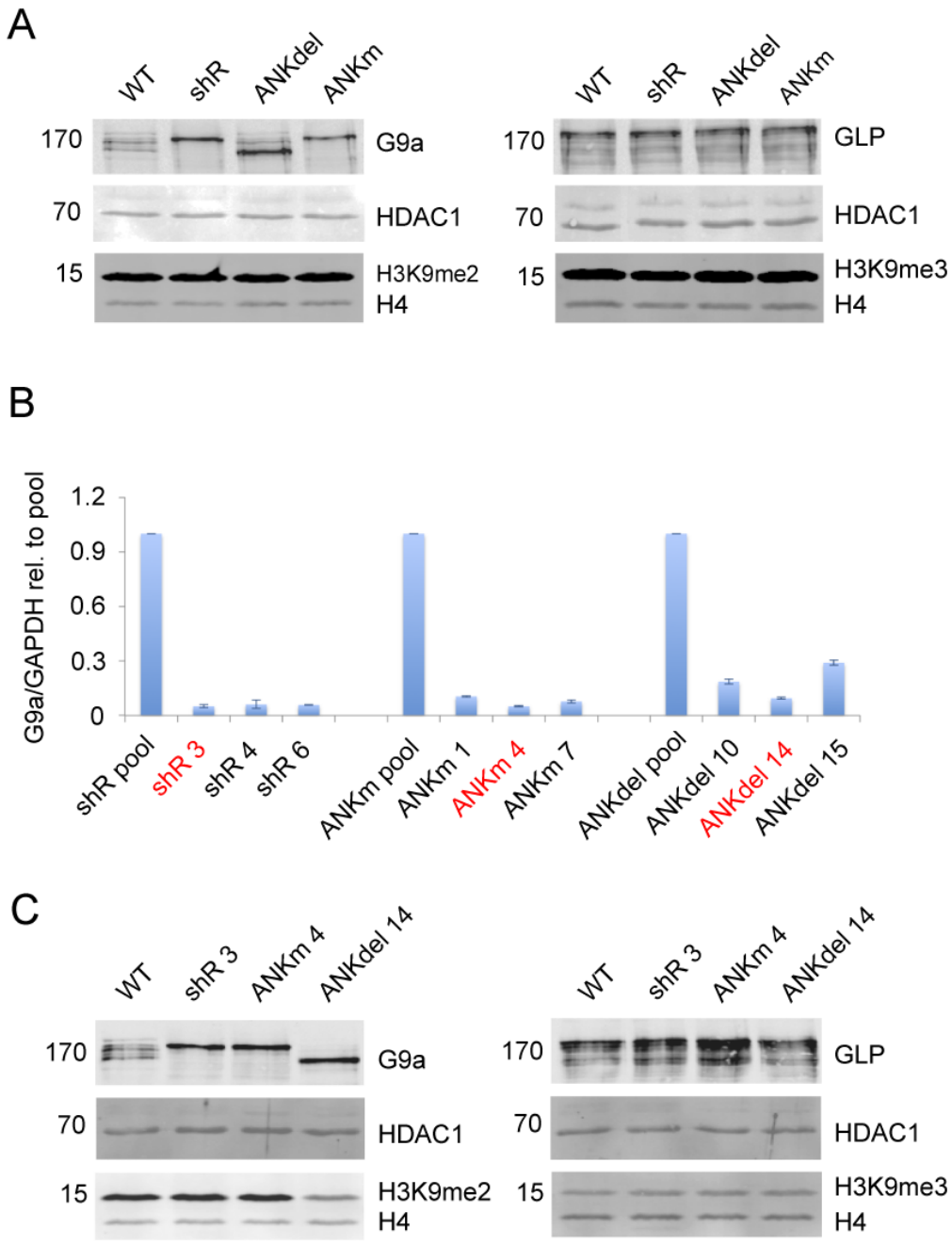


Figure 6.3 Generation of the ANKm and the ANKdel ES cell lines.

(**A**) Western blots detecting G9a, GLP and H3K9me levels in wild-type, shR, ANKm and ANKdel ES cells. (**B**) Quantitive PCR detects endogenous G9a mRNA levels in shR, ANKm, ANKdel pool clonal cell lines and their individual single clonal cell lines. The single clonal cell lines marked in red were used for the further analyses. (**C**) Western blots detecting G9a, GLP and H3K9me levels in shR, ANKm and ANKdel single clonal cell lines. shR represents shRNA resistant form of wild-type G9a; ANKm and ANKdel represent shRNA resistant forms of G9a containing either mutations in the ANK domain or the whole ANK domain deletion. HDAC1 and H4 are used as the loading controls for G9a and H3K9me, respectively.

These experiments showed that the levels of exogenous G9a and H3K9me3 in all these clonal cell lines were similar (Figure 6.3C). However, the levels of GLP and H3K9me2 varied among these ES cell lines. Compared with ANKdel ES cells, both shR and ANKm ES cells display higher levels of GLP and H3K9me2 (Figure 6.3C). However, compared with wild-type ES cells, ANKdel ES cells still contain similar levels of GLP, but less amount of H3K9me2 (Figure 6.3C). Taken together, after endogenous G9a was knocked down, all the exogenous forms of G9a can be stably expressed in ES cells. However, since only the level of H3K9me2 in ANKdel ES cells was impaired, different mutants within exogenous forms of G9a might exert distinct influences in function of G9a.

6.2.2 ANK domain mutant and deleted forms of G9a stably bind to chromatin

Before carrying out investigation of imprinted DNA methylation in shR, ANKm and ANKdel ES cells, I examined chromatin binding patterns of shRNA resistant wild-type, ANKm and ANKdel forms of G9a protein by step-wise salt extraction of nuclear proteins followed by Western blotting assays (Figure 6.4). These experiments showed that, in all these ES cell lines, the majority of histones dissociated from chromatin in 500 mM NaCl fraction, which indicated that step-wise salt extraction of nuclear proteins worked as expected. In addition, the levels of histone H4 were comparable among these cell lines in 500 mM NaCl fraction and the levels of H3K9me2 were consistent with my previous Western blotting results (Figure 5.5B). Furthermore, these experiments also showed that similar to wild-type G9a, both ANKm and ANKdel forms of G9a still tightly bind to chromatin in ES cells, since the majority of these mutant G9a proteins were

present in the fraction that extracted with benzonase and 500 mM NaCl (Figure 6.4). Taken together, these observations indicated that mutations that I generated have no effect on the chromatin loading patterns of G9a protein and also raised the question: will these mutant forms of G9a still play a role in the maintenance of imprinted DNA methylation in ES cells?

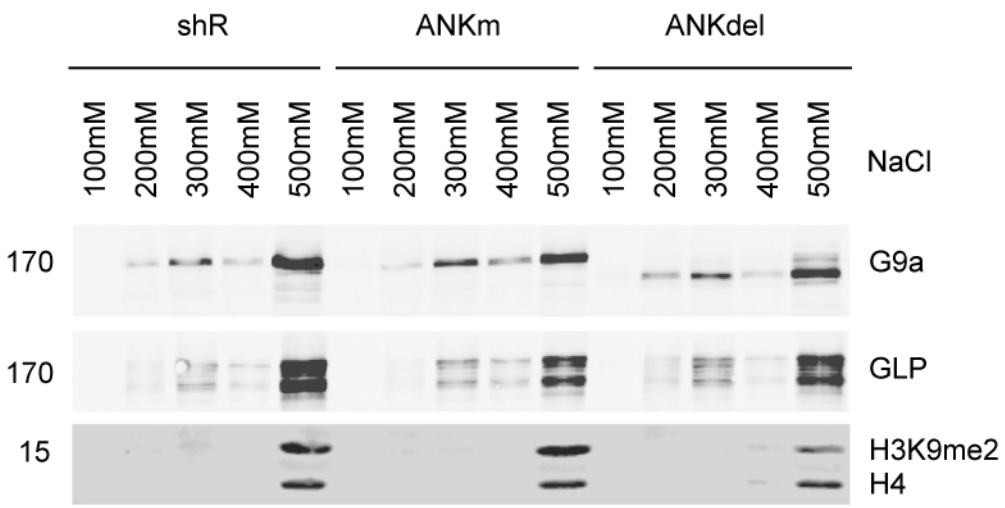


Figure 6.4 The majority of the ANKm and the ANKdel G9a stably bind to the chromatin.

Fractions of the step-wise extraction of the nuclear protein from shR, ANKm and ANKdel cell lines were analysed by Western blotting for G9a, GLP, H3K9me2 and H4. shR represents shRNA resistant form of wild-type G9a; ANKm and ANKdel represent shRNA resistant forms of G9a containing either mutations in ANK domain or the whole ANK domain deletion. H4 is a component of nuclesome and served as a control protein for the step-wise extraction of nuclear proteins.

6.2.3 Imprinted DNA methylation is stable in ES cells that expressing ANKm G9a

To investigate the patterns of imprinted DNA methylation in ES cells that express ANKm G9a, I first carried out bisulfite DNA sequencing in shR and ANKm ES cells. At *Igf2r* and *Snrpn* ICRs, the results in shR ES cells displayed normal patterns of imprinted DNA methylation (Figure 6.5). Interestingly, ANKm ES cells also display normal patterns of imprinted DNA methylation at these ICRs (Figure 6.5).

To validate bisulfite DNA sequencing results and further investigate the patterns of imprinted DNA methylation at more ICRs, I performed 5meC MeDIP in shR and ANKm ES cells at two paternal ICRs: *H19-Igf2* and *Rasgrf1* and eight maternal ICRs: *Igf2r*, *Snrpn*, *Zac1*, *Peg3*, *Peg10*, *Peg13*, *GnasXL* and *Kcnq1ot1*. For 5meC MeDIP, *Ndufa* and *Dpep3* again served as negative and positive control regions, respectively. Interestingly, compared with shR ES cells, seven ICRs displayed similar levels of imprinted DNA methylation in ANKm ES cells (Figure 6.6). However, three ICRs showed loss of imprinted DNA methylation in ANKm ES cells. This might due to less number of CpG sites involved in these DMRs and consequently, loss of DNA methylation at one or several CpG sites might influence the result of 5meC MeDIP. Taken together, these observations indicate that ANKm G9a, which is unable to bind to H3K9-methylated H3 tail (Collins et al., 2007), still plays a role in the maintenance of imprinted DNA methylation in ES cells.

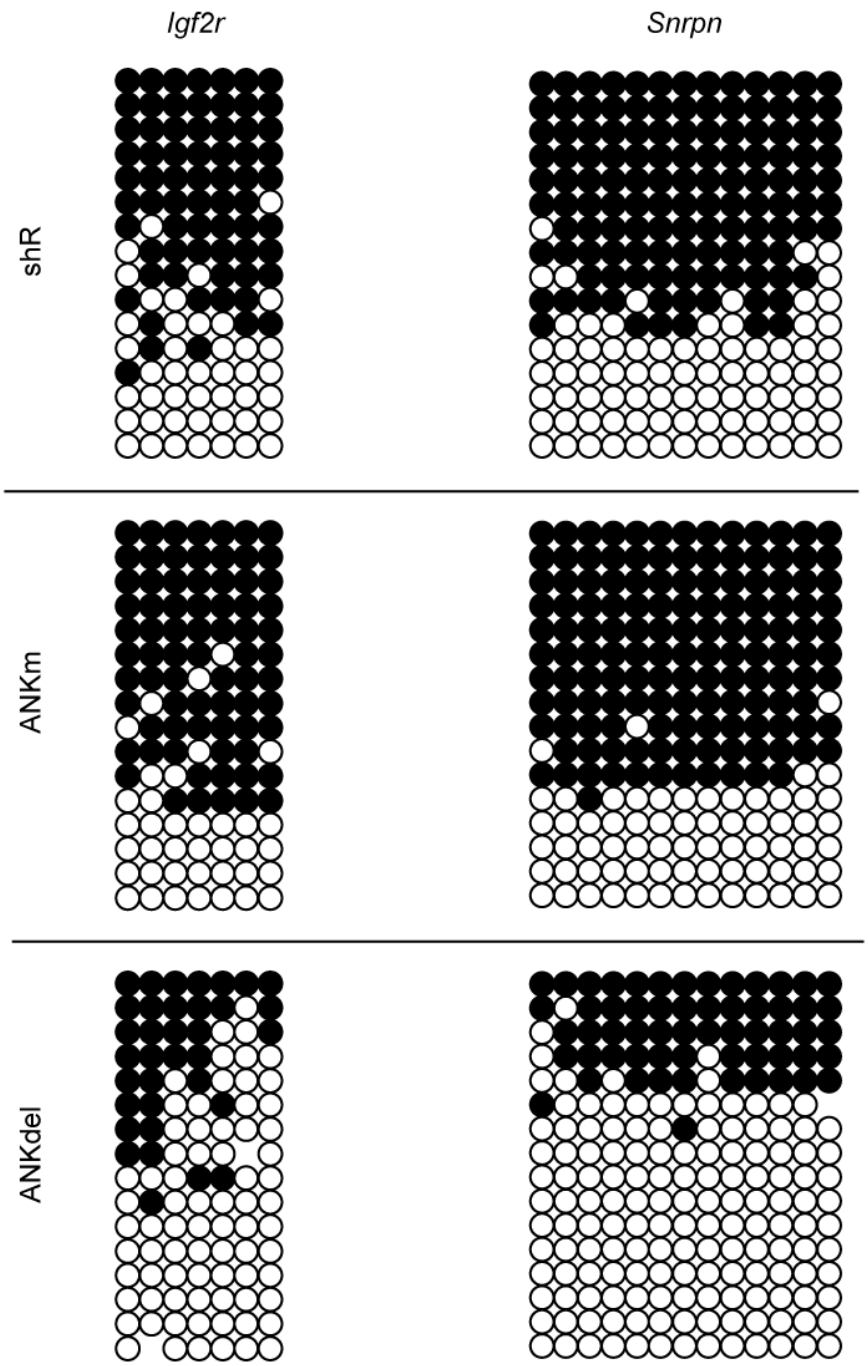


Figure 6.5 Patterns of DNA methylation at *Igf2r* and *Snrpn* ICRs in the ANKm and the ANKdel ES cells.

Both *Igf2r* and *Snrpn* are maternally derived ICRs, which display reduced levels of DNA methylation in ANKdel ES cells, but not in shR and ANKm ES cells. shR represents shRNA resistant form of wild-type G9a; ANKm and ANKdel represent shRNA resistant forms of G9a containing either mutations in ANK domain or the whole ANK domain deletion. Each black or white circle indicates an individual methylated or unmethylated CpG site, respectively. Each row represents a single DNA strand.

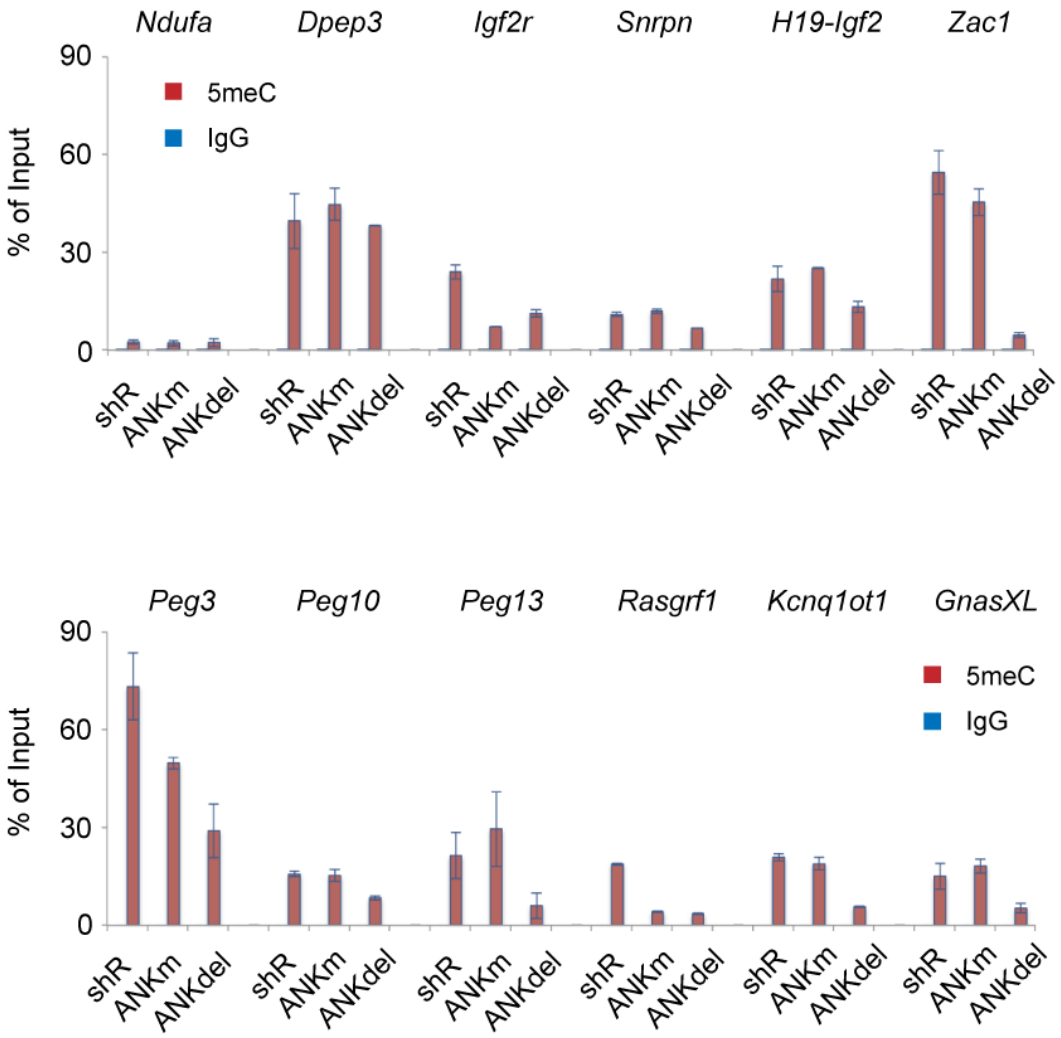


Figure 6.6 Imprinted DNA methylation is reduced in the ANKdel but not the ANKm ES cells.

Investigation of DNA methylation by 5mC MeDIP detects reduced methylation levels at the ICRs in the ANKm and the ANKdel ES cells in comparison with controls (shR). *Ndufa* and *Dpep3* served as the negative and positive control regions, respectively. *H19-Igf2* and *Rasgrf1* are two paternal ICRs, whereas *Igf2r*, *Snrpn*, *Zac1*, *Peg3*, *Peg10*, *Peg13*, *Kcnq1ot1* and *GnasXL* are the maternal ICRs. Among the detected ICRs, only three display hypomethylation in the ANKm ES cells, whereas all of them show hypomethylation in the ANKdel ES cells. shR represents shRNA resistant form of wild type G9a; ANKm and ANKdel represent shRNA resistant forms of G9a containing either mutations in ANK domain or the whole ANK domain deletion. The Y-axis value corresponds to percentage enrichment relative to input. IgG (blue bars) was used as the negative control for 5meC MeDIP. Error bars represent 2 times standard deviations of technical triplicates.

6.2.4 Imprinted DNA methylation is reduced in ES cells that expressing ANKdel G9a

I also investigated patterns of imprinted DNA methylation in ANKdel ES cells. I first performed bisulfite DNA sequencing and found that compared with shR ES cells, ANKdel ES cells displayed reduction of imprinted DNA methylation at *Igf2r* and *Snrpn* ICRs (Figure 6.5). Additionally, I also found that all the ICRs that investigated by 5meC MeDIP displayed hypo-methylation in ANKdel ES cells, including *Igf2r*, *Snrpn*, *H19-Igf2, Zac1*, *Peg3*, *Peg10*, *Peg13*, *GnasXL*, *Kcnq1ot1 and Rasgrf1* (Figure 6.6). These results indicate that ANK domain of G9a is essential for the maintenance of imprinted DNA methylation in ES cells and G9a is likely to protect ICRs from loss of DNA methylation via directly interacts with DNMTs as reported previously (Esteve et al., 2006; Epsztejn-Litman et al., 2008).

6.2.5 Both ANKm and ANKdel forms of G9a interact with DNMTs

To investigate whether ANKm or ANKdel forms of G9a interact with DNMTs, I carried out co-immunoprecipitation (Co-IPs) experiments with antibodies against G9a and DNMTs using nuclear extracts that prepared from shR, ANKm and ANKdel ES cells. The results showed that in all these ES cells, antibodies against G9a could pull down GLP and *vice versa* (Figure 6.7A). This observation indicates that mutations within G9a protein had no effect on dimerization between G9a and GLP *in vivo*.

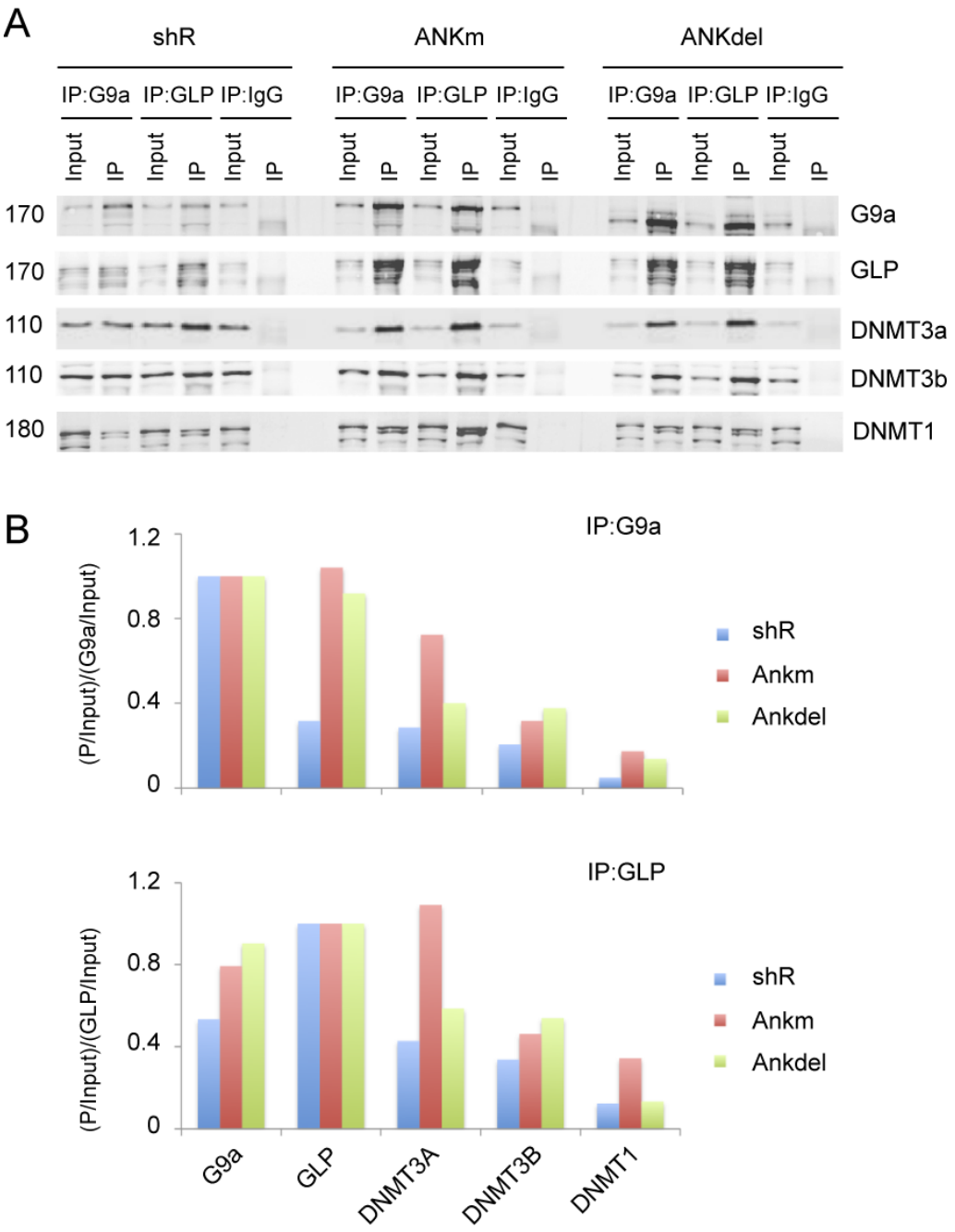


Figure 6.7 Reduced amount of DNMT3A and DNMT1 interact with the G9a/GLP complex in the ANKdel ES cells.

(A) Co-immunoprecipitation assays confirmed the interaction between either G9a or GLP and DNMTs in shR, ANKm and ANKdel ES cells. Total nucler protein were extracted from shR, ANKm and ANKdel ES cells and Co-IP assays were performed by using G9a, GLP, DNMT3A, DNMT3B and DNMT1 antibodies againest the pulled down nuclear protein by either G9a or GLP antibodies. Mouse IgG antibody was used as a control antibody for G9a and GLP. shR represents shRNA resistant form of wild-type G9a; ANKm and ANKdel represent shRNA resistant forms of G9a containing either mutations in ANK domain or the whole ANK domain deletion. (B) Quantification of the Co-IP asssys indicated that compared with the ANKm-G9a/GLP complex, there was less amount of DNMT3A and DNMT1 that interact with the ANKdel-G9a/GLP complex.

Moreover, Co-IP experiments also showed that both ANKm and ANKdel forms of G9a could interact with DNMTs, including DNMT1, DNMT3A and DNMT3B (Figure 6.7A). Additionally, the ratio of DNMTs/G9a showed that compared with wild-type (shR) G9a, there were similar amounts of DNMT3B interacted with either ANKm or ANKdel forms of G9a (Figure 6.7B). By contrast, the amount of DNMT1 and DNMT3A that interacted with ANKdel form of G9a is less than the amount that interacted with shR and ANKm forms of G9a (Figure 6.7B). These observations indicated that interaction between G9a and DNMT3A or DNMT1 might play an important role in the maintenance of imprinted DNA methylation in ES cells.

6.3 Discussion

After removal of endogenous *G9a* in ANKm ES cells, I found that among ten investigated ICRs, seven of them still contain normal patterns of imprinted DNA methylation (Figure 6.6). This observation indicates that abolishing of interaction between ANK domain of G9a protein and H3K9-methylated histone H3 tail does not affect imprinted DNA methylation that is present at most of ICRs in ES cells. Since compared with wild-type ES cells, there is more GLP in ANKm ES cells (Figure 6.3C), it is possible that imprinted DNA methylation is maintained by GLP/GLP homodimers. Alternatively, since ANKm G9a still forms a dimer with GLP and the intact ANK domain within GLP protein could recognise H3K9-methylated histone H3 tail and bind to chromatin, ANKm G9a might indirectly bind to ICRs and maintain imprinted DNA methylation by recruitment of DNMTs to ICRs via the ANK domain of GLP.

Actually, previous studies have shown that apart from ANK domain, SET domain of G9a could also recognise and bind to unmethylated or H3K9me1 modified histone H3 tail (Collins et al., 2008). Since there is no reduction of H3K9me1 in ANKm ES cells, the

interaction between SET domain of G9a and H3K9me1 might play a role in the recruitment of ANKm G9a to chromatin. However, compared with ANK domain of G9a, the interaction between SET domain of G9a and histone H3 tail is more unstable and transient (Collins et al., 2007). In addition, since GLP displays similar chromatin binding property to G9a (Collins et al., 2008) and ANKm G9a still interacts with GLP, the chromatin binding of ANKm G9a might occur indirectly through the interaction between G9a partner protein GLP and H3K9me1/2. Moreover, since G9a/GLP/WIZ tri-molecular complex is the predominant form of G9a that exist *in vivo* (Tachibana et al., 2005; 2008 Ueda et al., 2006) and C2H2 zinc finger WIZ might function as a potential DNA-binding protein, the chromatin loading of ANKm G9a might also depend on the interaction between WIZ and DNA.

To investigate how ANKm G9a binds to chromatin, amino acids within the pre-SET domain of G9a, which are essential for dimerization between G9a and GLP, would need to be mutated in ANKm G9a. Step-wise salt extraction of nuclear proteins combined with Western blotting assays would then have to be carried out to investigate whether ANKm G9a binds to chromatin depending on its own SET domain or the other functional domain(s) within G9a/GLP complex.

In this chapter, I also investigated the patterns of imprinted DNA methylation in ANKdel ES cells and found that all the investigated ICRs reduced the levels of imprinted DNA methylation (Figure 6.6). This observation indicated that ANK domain of G9a is essential for the maintenance of imprinted DNA methylation in ES cells. However, step-wise salt extraction of nuclear proteins followed by Western blottings assays as well as Co-IP experiments using antibodies against G9a and DNMTs indicate that ANKdel G9a is still able to bind to chromatin and interacts with DNMTs (Figure 6.4 and 6.7A).

However, the levels of DNMT3A and DNMT1 that interact with ANKdel G9a are less than the levels that interact with shR and ANKm G9a. These observations indicated that G9a is likely to maintain imprinted DNA methylation in ES cells via its ANK domain directly recruits of DNMT3A or DNMT1 to ICRs.

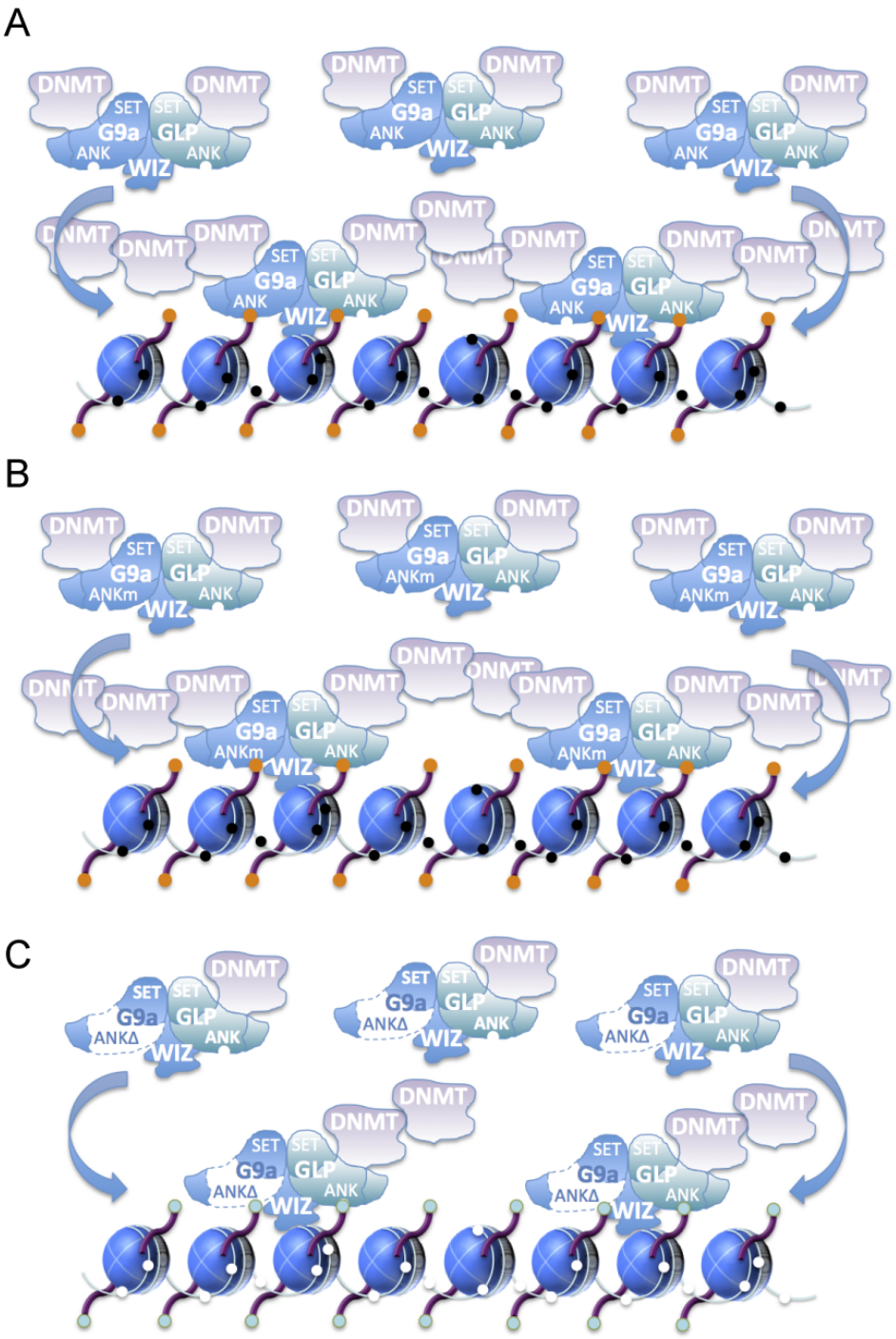


Figure 6.8 Model of maintenance of the imprinted DNA methylation by G9a/GLP complex.

(A) In wild-type ES cells, G9a/GLP complex binds to the ICRs and maintains imprinted DNA methylation by recriuting DNMTs via their ANK domains. (B) In ANKm ES cells, G9a/GLP complex binds to the ICRs via the ANK domain of G9a or WIZ, and maintains imprinted DNA methylation by recriuting DNMTs via the ANK domain of GLP. (C) In ANKdel ES cells, G9a/GLP complex binds to the ICRs via the ANK domain of GLP or WIZ. However, without the ANK domain of G9a, the local concentration of DNMTs is insufficient for the maintenance of imprinted DNA methylation. The black and white dots represent for the methylated and unmethylated cytosines within the ICRs, respectively. The oragne dots represent for the H3K9me2, whereas the green dots represent for the H3K9me0/1.

Since previous studies have shown that both ANK domain and SET domain within G9a or GLP could interact with histone H3 tail (Collins et al., 2008) and either G9a or GLP could interact with DNMTs through their ANK domains (Epsztejn-Litman et al., 2008), in theory, G9a/GLP complex could bind to ICRs through one of the ANK domains and meanwhile interacts with DNMTs via the other ANK domain in wild-type ES cells (Figure 6.8). However, in ANKdel ES cells, since G9a-ANKdel/GLP complex contains only one ANK domain, the only possibly for G9a-ANKdel/GLP complex binds to chromatin is via the interaction between ANK domain of GLP and H3K9me1. In addition, this hypothesis could also explain why imprinted DNA methylation is reduced in ANKdel ES cells but not in ANKm ones.

Chapter 7 Dimerization between G9a with GLP is essential for the maintenance of the imprinted DNA methylation in ES cells

7.1 Introduction

G9a/GLP complex serves as the predominant form of either G9a or GLP that exist in ES cells (Tachibana et al., 2005; Ueda et al., 2006). In addition, re-expression of either mutant G9a or GLP in corresponding knockout ES cells identified that NHLC (1165-1168) and NHHC (1198-1201) motifs within pre-SET domain of G9a and GLP respectively, are essential for dimerization between G9a and GLP (Tachibana et al., 2008). Moreover, it has also been found that replacement of NH (1165-1166) by LE in G9a protein is sufficient to abolish more than 95% of dimerization between G9a and GLP in ES cells (Tachibana et al., 2008).

In my previous chapter, I have shown that ANKm G9a still could maintain imprinted DNA methylation at many ICRs, which raised the question that how does ANKm G9a protect ICRs from loss of DNA methylation in ES cells. In previous chapter, I also carried out co-immunoprecipitation experiments using antibodies against G9a and GLP and found that ANKm G9a is still able to form a dimer with GLP. Since GLP also contains an ANK domain and displays similar chromatin loading patterns to those of G9a, it is likely that ANK domain of GLP plays a key role for ANKm G9a in the maintenance of imprinted DNA methylation in ES cells. In addition, my previous observations also indicate that dimerization between G9a and GLP might be required for the maintenance of imprinted DNA methylation in ES cells.

To further study how ANKm G9a maintain imprinted DNA methylation and whether dimerization between G9a and GLP is essential for the maintenance of imprinted DNA methylation in ES cells, I aimed to disrupt of dimerization between G9a and GLP and investigate the role of dimerization mutant G9a at imprinted loci without effects of GLP.

7.2 Results

7.2.1 The dimerization deficient forms of G9a are unstable in the ES cells

Since deletion of whole NHLC motif within pre-SET domain of G9a protein might affect structure of the SET domain, to disrupt dimerization between G9a and GLP, I replaced NH that within dimerization association motif by LE in G9a protein (Figure 7.1). Together with nucleotide mutations that used for generation of shR, ANKm and ANKdel forms of G9a, all created mutants within G9a cDNA were confirmed by DNA sequencing (Figure 7.2).

Next, since my previous data have shown that neither wild-type nor catalytic inactive G9a can restore loss of imprinted DNA methylation in *G9a-/-* ES cells (see chapter 3), to investigate the role of dimerization-deficient G9a on imprinted DNA methylation, I initiated studies by using wild-type ES cells. I first linearized dimerization-deficient G9a expression plasmids and transfected these plasmids into wild-type ES cells. Then, I knocked down endogenous *G9a* by shRNA and picked four single clonal ES cell lines in each of dimerization-deficient ANKm and ANKdel G9a (dmANKm and dmANKdel) pool clonal ES cells (see Materials and Methods) (Figure 7.3A).

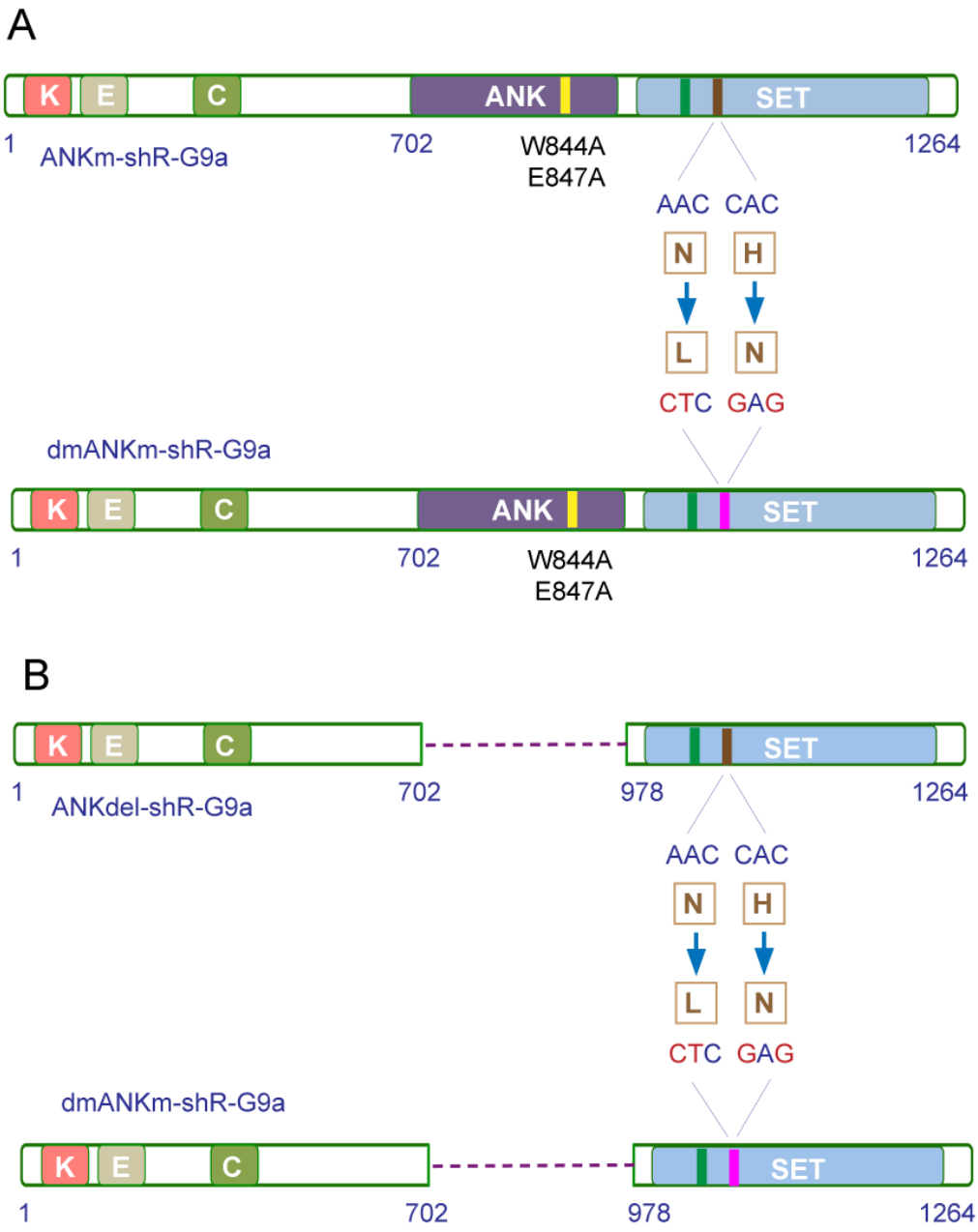


Figure 7.1 Dimerisation mutations introduced into ANKm and ANKdel forms of G9a cDNAs and protein.

(A) Schematic diagram depicting the generation of dimerisation-deficient ANK domain mutant (ANKm) shRNA resistant (shR) form of G9a. (B) Generation of the dimerisation-deficient ANK domain deleted (dmANKdel) sh RNA resistant (shR) form of G9a. See Figure X (Chapter 6) for the generation of the ANKm shR and ANKdel shR G9a. The green bar represents the nucleotide mutations that render G9a RNA resistant to shRNA without altering the amino acid sequence of G9a protein. The yellow bar represents the W844A and E847A point mutations within the ANK domain of G9a protein, which disrupts binding to H3K9me1 and H3K9me2. The change from brown bar to pink represents the N1165L and H1166N point mutations within the dimerization motif of G9a protein. K is a lysine rich region; E is a glutamic acid-rich region; C is a cysteine-rich region; ANK represents the ankyrin repeats domian; SET represents the SET domain.

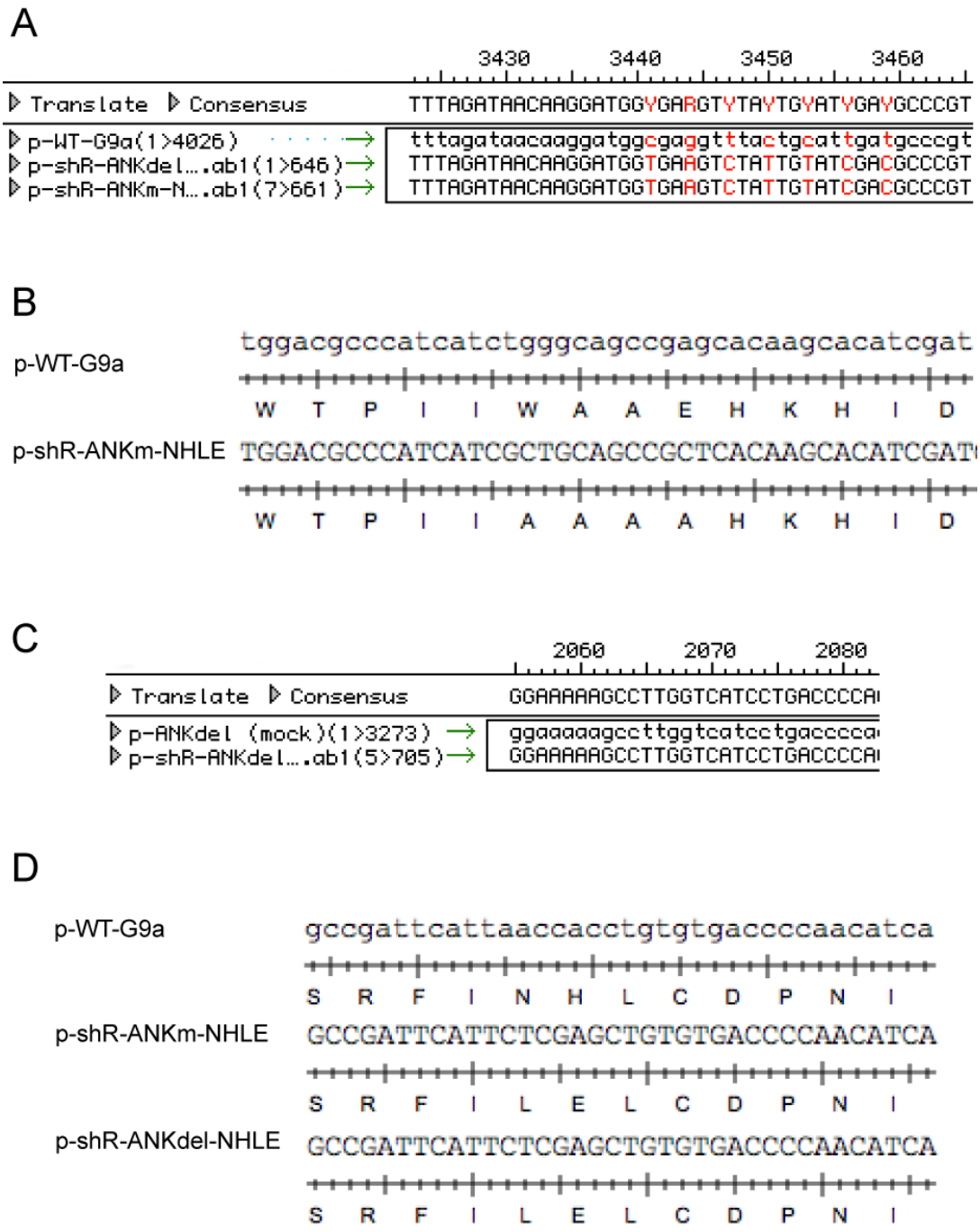


Figure 7.2 Sequencing confirms the dimerization-disrupting mutations that introduced into G9a cDNA.

(A) Sequencing comfirms the short hairpin RNA (shRNA) resistant mutations that introduced into dimerization mutant forms of G9a cDNA. (B) Sequencing confirms the presence of W844A and E847A mutations that encoded by the dmANKm form of G9a cDNA. (C) Sequencing confirms the deletion of the entire ANK domain that encoded by the dmANKdel form of G9a cDNA. p-ANKdel (mock) represents for the ANKdel form of G9a cDNA, which is virtually generated by the DNASTAR software. (D) Sequencing confirms the presence of N1165L and H1166E mutations in the G9a/GLP dimerization motif that encoded by the mutant forms of G9a cDNA.

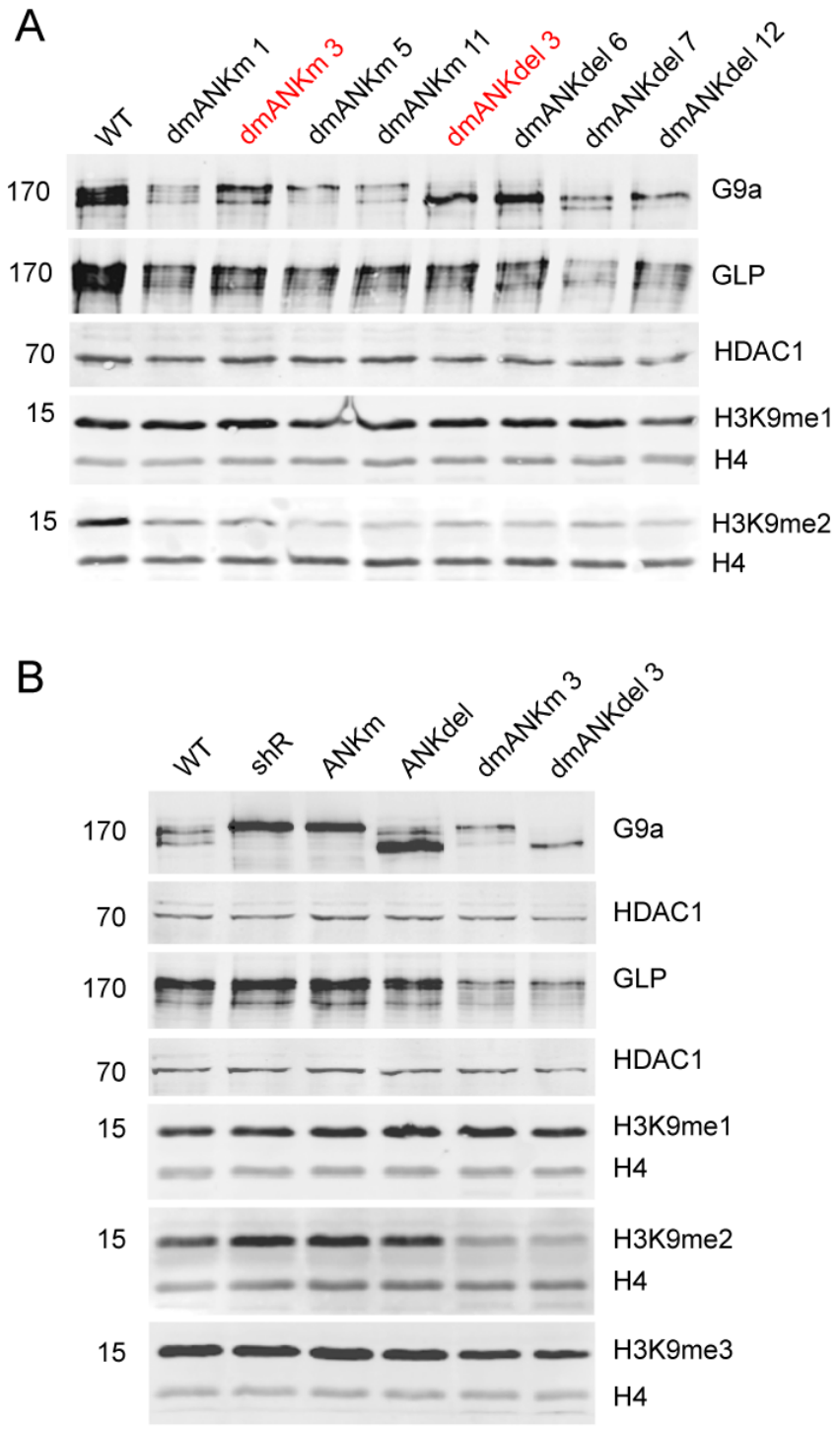


Figure 7.3 Generation of the dmANKm and the dmANKdel ES cell lines.

(A) Western blots detecting G9a, GLP and H3K9me levels in wild-type and several clonal ES cell lines that expressing dimerization-deficient ANK domain mutated (dmANKm) and dimerization-deficient ANKdomain deleted (dmANKdel) G9a. The single clonal cell lines marked in red were used for the further analyses.(B) Western blots detecting G9a, GLP and H3K9me in dmANKm and dmANKdel ES cells. See Chapter 6 for generation of the shR, ANKm and ANKdel ES cell lines. Antibodies against HDAC1 and H4 serve as the loading controls for G9a or GLP and H3K9me.

I then carried out Western blotting assays on wild-type and all single clonal ES cell lines that expressing either dmANKm or dmANKdel forms of G9a. The results showed that there was less G9a protein in the clonal ES cell lines that expressing dimerization-deficient form of G9a. This is consistent with published observations (Tachibana et al, 2008) and together suggests that without forming a dimer with GLP, G9a monomer cannot exist stably *in vivo*. To select the ideal dmANKm and dmANKdel single clonal ES cell lines, which express comparable levels of exogenous G9a, I chose dmANKm-3 and dmANKdel-3 for downstream experiments.

7.2.2 The levels of GLP and H3K9me2 were reduced in the dmANKm and dmANKdel ES cells

In chapter 6, I showed that compared with wild-type ES cells, ANKm ES cells contain normal levels of GLP and H3K9me2 (Figure 6.3C). However, in dmANKm ES cells, Western blotting results showed that both the levels of GLP and H3K9me2 are reduced (Figure 7.3B). These observations suggest that disruption of dimerization between G9a and GLP not only affects stability of G9a, but also destabilises GLP. These observations are consistent with previous studies, which conclude that G9a/GLP complex is the predominant form of either G9a or GLP *in vivo* (Tachibana et al., 2002, 2005). In addition, these observations also suggest that reduced levels of G9a and GLP are not sufficient for the maintenance of H3K9me2 in the ES cells, which is consistent with that G9a and GLP function as main writers for the establishment and maintenance of H3K9me2 in mammalian cells (Tachibana et al., 2005; Ueda et al., 2006).

Interestingly, the global level of H3K9me3 in dmANKm and dmANKdel ES cells displayed slight reduction when compared to wild-type ES cells, while the levels of H3K9me1 remained unchanged (Figure 7.3B). These observations suggest that

G9a/GLP-mediated H3K9me2 is required as the partial subtracts for histone modifiers of H3K9me3 in ES cells. In addition, these observations and my previous data (Figure 3.4A, 4.1C and 4.7B) are conflicted with previous studies (Tachibana et al., 2002; Peters et al., 2003; Rice et al., 2003) and indicate that apart from G9a and GLP, some other histone modifiers must play a role in the regulation of level of H3K9me1 in ES cells.

7.2.3 Dimerization-deficient forms of G9a stably bind to the chromatin

My previous experiments described in Chapter 6 have suggested that G9a is likely to protect ICRs from loss of imprinted DNA methylation by directly recruitment of DNMTs via its ANK domain. Therefore, I first investigated whether dimerization-deficient forms of G9a could bind to chromatin by performing step-wise extraction of nuclear proteins combined with Western blotting assays (Figure 7.4).

LSH is a chromatin remodelling helicase (Myant et al., 2011), which displayed step-wise dissociation from chromatin at NaCl concentrations higher than 200 mM. The patterns of LSH dissociation from chromatin in shR, dmANKm and dmANKdel ES cells were identical (Figure 7.4). STELLA is another chromatin binding protein and its chromatin loading depends on the presence of H3K9me2 (Nakamura et al., 2007, 2012). Western blotting for STELLA showed that in shR ES cells, the majority of STELLA dissociated from the chromatin at 200-300 mM NaCl concentration, whereas in dmANKm and dmANKdel ES cells, some of STELLA protein dissociated from chromatin at 0 mM salt concentration. This observation is consistent with reduced levels of H3K9me2 in these ES cells (Figure 7.3). In addition, RAN is another nuclear protein, which is unable to bind to chromatin and plays a role in nuclear transport through nuclear pores. The results showed that the majority of RAN is extracted from nuclei at low salt

concentration, which is as expected (Figure 7.4). Moreover, Western blottings for histones H3 and H4 indicated, as expected, that histones did not dissociate from chromatin at salt concentration below 500 mM (Figure 7.4). This is consistent with that histones serve as core components of chromatin. Taken together, Western blotting results of all the above control proteins, including LSH, STELLA, RAN and histones, indicated that step-wise extraction of nuclear proteins that performed in these ES cells worked as expected.

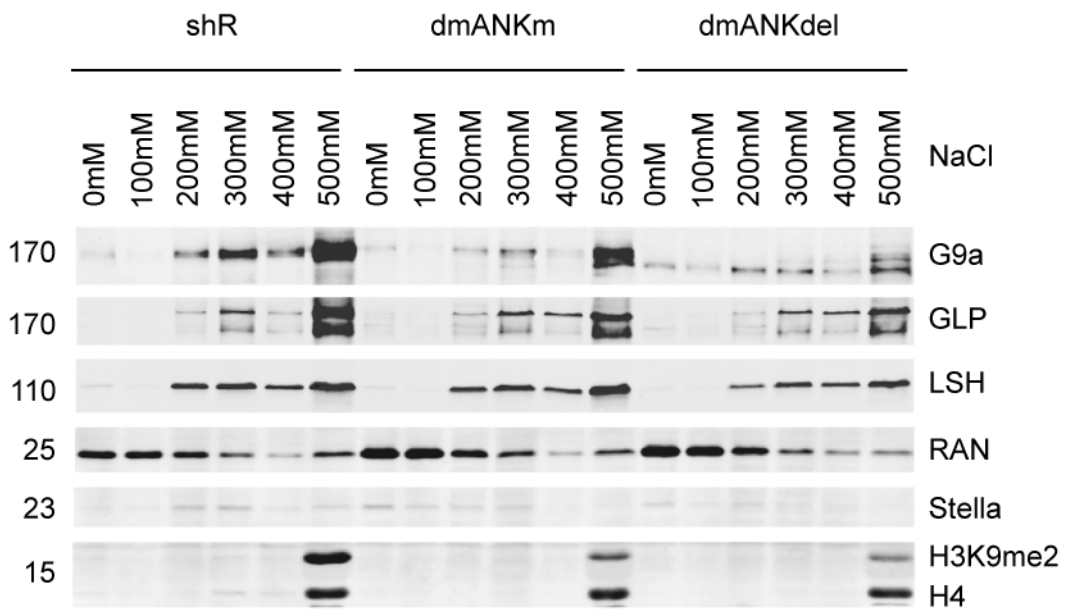


Figure 7.4 Dimerization-deficient forms of G9a stably binds to the chromatin

Fractions of the step-wise extraction of nuclear protein from ES cells that expressing shR, dmANKm and dmANKdel forms of G9a were analysed by Western blotting for G9a, GLP, H3K9me2 and H4. In these ES cells, the majority of G9a and GLP dissociates from chromatin at 300-500 mM NaCl concentration. Chromatin remodelling protein LSH dissociates from chromatin at 200 mM NaCl concentration. The mojority of nuclear export protein RAN, which does not bind to chromatin, is extracted at low salt concentration (0-100 mM NaCl). The global levels of H3K9me2 are reduced in dmANKm and dmANKdel ES cells. Consistently, the H3K9me2 binding protein STELLA dissociates from chromatin at 200 mM NaCl concentration in control shR ES cells, but in dmANKm and dmANKdel ES cells, STELLA dissociates at 0 mM NaCl concentration. H4 is a component of nucleosomes and serves as a control protein for the step-wise extraction of nuclear proteins.

Strikingly, Western blotting results indicate that both dmANKm and dmANKdel G9a as well as GLP are still tightly bound to chromatin, since the majority of dimerization-deficient G9a and GLP dissociated from chromatin at 300-500 mM NaCl concentration (Figure 7.4). In addition, these observations indicate that monomeric forms of G9a and GLP are still able to bind to chromatin. In the case of G9a, such binding can even occur in the absence of ANK domain.

7.2.4 Dimerization-deficient forms of G9a do not interact with the GLP and DNMTs

To investigate whether dmANKm and dmANKdel forms of G9a have lost ability to form a dimer with GLP, I carried out co-immunoprecipitation experiments (Co-IPs) using antibodies against G9a and GLP in shR, dmANKm and ANKdel ES cells. The results showed that in dmANKm and dmANKdel ES cells, antibodies against G9a could not pull down GLP and *vice versa* (Figure 7.5). These data confirm that dmANKm and dmANKdel forms of G9a are indeed unable to interact with GLP.

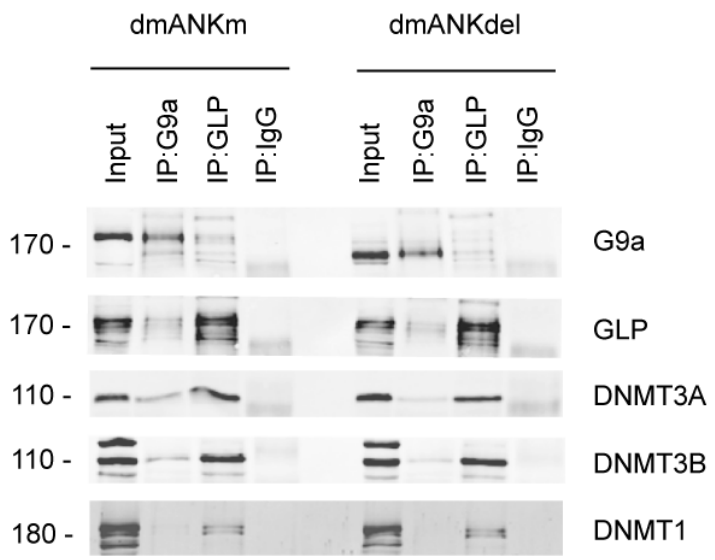


Figure 7.5 Both the dmANKm and the dmANKdel G9a lost the interaction with DNMTs.

Co-immunoprecipitation (Co-IP) assays with antibodies against G9a and GLP confirmed that the dimerization-deficient forms of G9a (dmANKm and dmANKdel) lost the interaction with GLP and *vice versa*. In addition, both the dmANKm and the dmANKdel G9a lost the interaction with DNMTs. Mouse IgG was used as a control for the non-specific antibody of the Co-IP experiments. The input lane corresponds to 1/10 of the total amount of nuclear extract used in the IPs. The Western blottings with resolved immunoprecipitated proteins were detected with antibodies shown on the right (G9a, GLP, DNMT3A, DNMT3B and DNMT1). Note that GLP retains its interaction with DNMTs in the absence of interaction with G9a.

.

In addition, to investigate whether monomeric forms of G9a and GLP could interact with DNMTs, I also carried out Co-IP experiments using antibodies against G9a, GLP and DNMTs. These experiments showed that in dmANKm and dmANKdel ES cells, GLP is still able to interact with DNMTs, whereas dmANKm and dmANKdel forms of G9a dramatically lost their interactions with DNMTs (Figure 7.5). These observations indicated that GLP does not require G9a to interact with DNMTs in either dmANKm or dmANKdel ES cells (Figure 7.5). Taken together, these observations clearly showed that without forming a dimer with GLP, both ANKm and the ANKdel forms of G9a are unable to interact with DNMTs in ES cells.

7.2.5 Imprinted DNA methylation is reduced in dmANKm ES cells

I next investigated the patterns of imprinted DNA methylation in dmANKm ES cells by performing bisulfite DNA sequencing. At all the investigated ICRs, *Igf2r*, *Snrpn* and *H19-Igf2*, the levels of the imprinted DNA methylation were reduced when compared to the levels in ANKm ES cells (Figure 7.6 and Figure 6.5, 6.6). In addition, to validate bisulfite DNA sequencing results and further investigate the patterns of imprinted DNA methylation at more ICRs, I carried out 5meC MeDIP at *Igf2r*, *Snrpn*, *H19-Igf2*, *Zac1*, *Peg3*, *Peg10*, *Peg13*, *GnasXL*, *Rasgrf1* and *Kcnq1ot1* imprinted loci in shR, dmANKm and dmANKdel ES cells. Compared with shR ES cells, all the investigated ICRs displayed hypo-methylation in dmANKdel ES cells (Figure 7.7). This observation is consistent with previous 5meC MeDIP results that carried out in ANKdel ES cells (Figure 6.6).

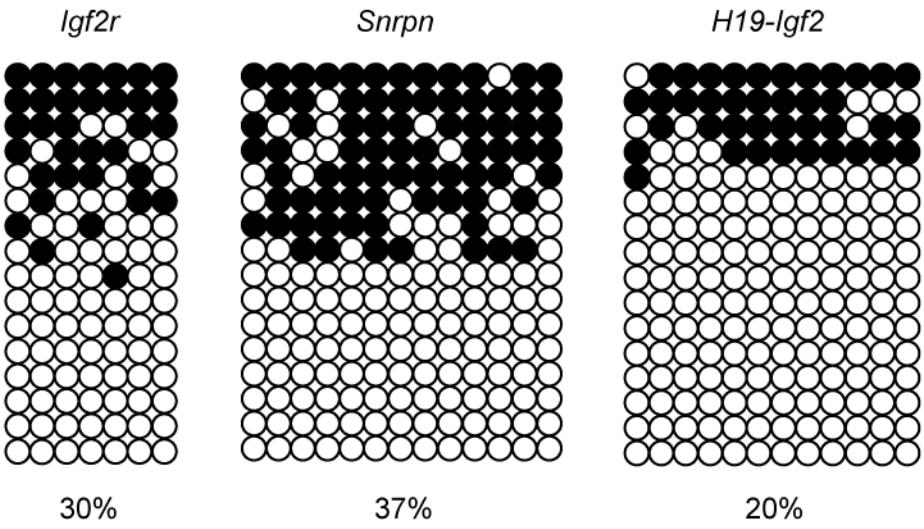


Figure 7.6 Patterns of DNA methylation at *Igf2r, Snrpn* and *H19-Igf2* ICRs in dmANKm ES cells.

Igf2r and *Snrpn* are maternally methylted ICRs, whereas *H19-Igf2* is a paternally methylated ICR. All these three ICRs displayed reduced DNA methylation levels in the dmANKm ES cells. Each black or white circle indicates an individual either methylated or unmethylated CpG site, respectively. Each row represents a single DNA strand.

Strikingly, unlike in ANKm ES cells (Figure 6.6), nearly all the investigated ICRs, except *Snrpn, Peg13* and *GnasXL*, displayed hypo-methylation in dmANKm ES cells (Figure 7.7). This observation suggests that without forming a dimer with GLP, dmANKm G9a cannot efficiently maintain imprinted DNA methylation in ES cells. Furthermore, this observation also confirms my previous hypothesis, which indicates that in ANKm ES cells, the patterns of imprinted DNA methylation is likely to be protected by GLP-mediated recruitment of DNMTs to ICRs.

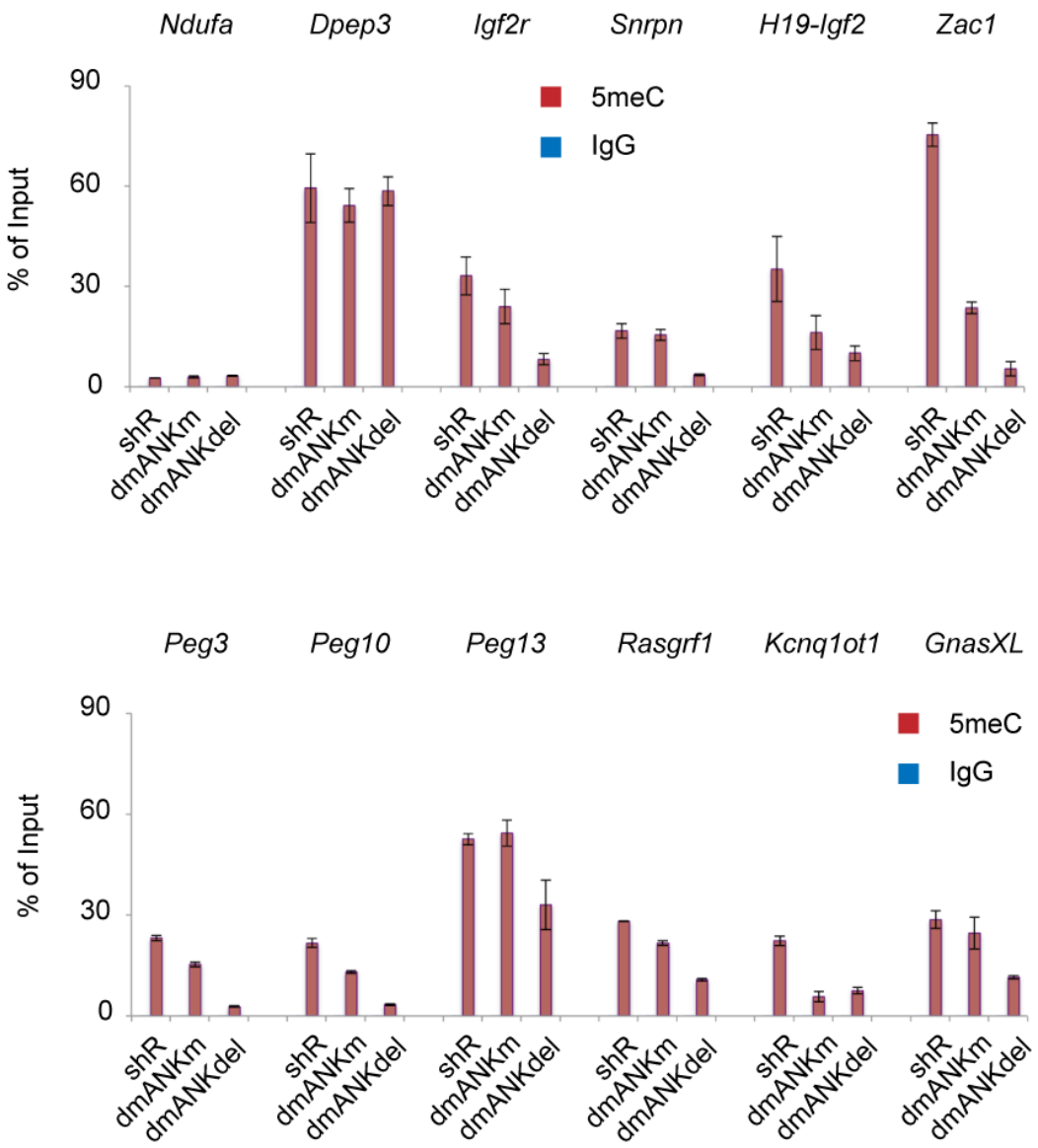


Figure 7.7 DNA methylation is reduced at the ICRs in the dmANKm and the dmANKdel ES cells.

Investigation of DNA methylation by 5mC MeDIP detects reduced methylation levels at the ICRs in the dimerisation-deficient ES cells (dmANKm and dmANKdel) in comparison with controls (shR). *H19-Igf2* and *Rasgrf1* are paternally methylated ICRs, whereas *Igf2r*, *Snrpn*, *Zac1*, *Peg3*, *Peg10*, *Peg13*, *Kcnq1ot1* and *GnasXL* are maternally methylated ICRs. Among the detected ICRs, seven display hypomethylation in the dmANKm ES cells, whereas all of them show hypomethylation in the dmANKdel ES cells. *Ndufa* and *Dpep3* serve as the negative and positive control regions, respectively.The Y-axis value corresponds to percentage enrichment relative to input. IgG (blue bars) was used as the negative control for 5meC MeDIP. Error bars represent 2 times standard deviations of technical triplicates.

7.3 Discussion

In this Chapter, I described experiments aiming to investigate how does ANKm G9a maintain imprinted DNA methylation in ES cells. Since previous co-immunoprecipitation assays have shown that ANKm G9a still forms a dimer with GLP and through ANK domain of GLP, ANKm-G9a/GLP complex could also interact with DNMTs, it is likely that ANKm G9a maintains imprinted DNA methylation by recruitment of DNMTs to ICRs via the ANK domain of GLP.

To eliminate the role of GLP in the maintenance of imprinted DNA methylation in ANKm ES cells, I generated dmANKm ES cell line (Figure 7.3). Step-wise salt extraction of nuclear proteins followed by Western blottings showed that without forming a dimer with GLP, dimerization-deficient ANKm (dmANKm) G9a are still tightly bound to chromatin in ES cells (Figure 7.4). Published data have reported that the catalytic SET domain of G9a can interact with H3K9me0/1 to perform its methyltransferase function (Collins et al., 2008). However, these interactions are likely to be short-lived and therefore it seems that interaction between ANK domain and H3K9me0/1 would stabilise the binding of G9a to chromatin. In addition, since dmANKm G9a should not be able to bind to either H3K9me1 or H3K9me2 in ES cells, it is unclear that how dmANKm G9a binds to chromatin. Since G9a/GLP complex contains a third stable component, the C2H2 zinc-finger protein WIZ, it is likely that WIZ might bind to DNA directly and plays a role in the recruitment of monomeric G9a to chromatin. It would therefore be important to investigate whether dmANKm G9a still interacts with WIZ and whether zinc-fingers protein WIZ binds to DNA, especially in a sequence-specific context.

To investigate the patterns of imprinted DNA methylation in dmANKm ES cells, I carried out bisulfite DNA sequencing and 5meC MeDIP and found that the levels of

imprinted DNA methylation is reduced in dmANKm ES cells (Figure 7.6 and 7.7). Compared with ANKm ES cells, where the majority of the ICRs carry normal patterns of the imprinted DNA methylation, loss of the imprinted DNA methylation in dmANKm ES cells suggests that without forming a dimer with GLP, dmANKm G9a is unable to maintain DNA methylation that present at ICRs. In addition, this observation also suggests that the stable patterns of imprinted DNA methylation in ANKm ES cells is likely to be maintained by ANKm G9a interaction with GLP, which could recruit DNMTs to ICRs via its ANK domain (Figure 7.8).

Taken together, G9a/GLP complex is likely to maintain imprinted DNA methylation in ES cells via binding to DNA and/or chromatin at ICRs mediated by WIZ or ANK domain-dependent recognition of H3K9me2, while simultaneously recruit DNMTs by ANK domain(s) within G9a, GLP or both proteins (Figure 7.8). Whether a single ANK domain can bind to chromatin and at the same time recruit DNMTs to ICRs is yet to be determined. In addition, it is also possible that G9a/GLP complex indirectly binds to ICRs via the interactions with other yet unknown chromatin binding proteins. However, this must occur only on the allele that carries DNA methylation.

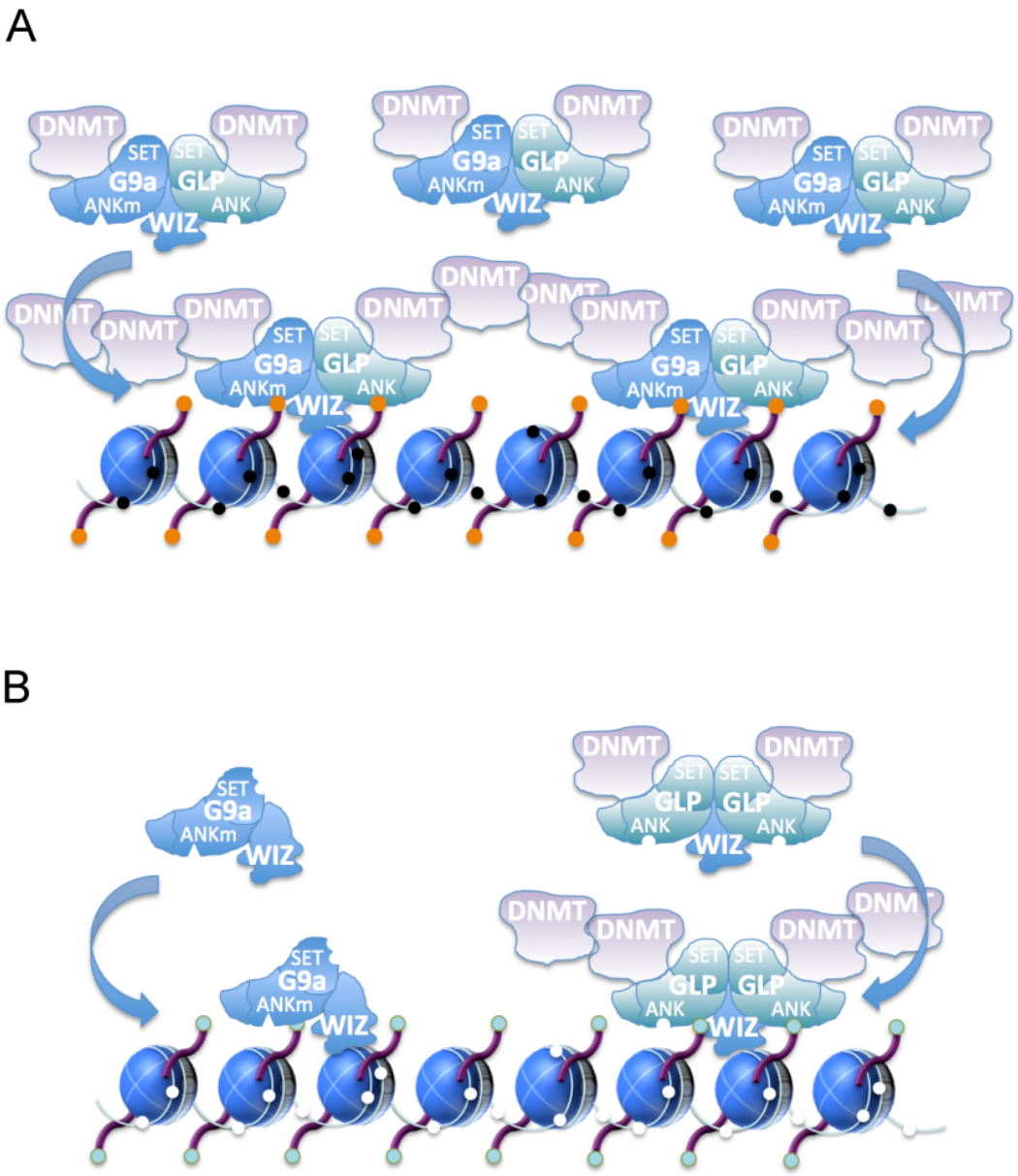


Figure 7.8 Model of maintenance of the imprinted DNA methylation by G9a/GLP complex.

(A) In ANKm ES cells, G9a/GLP complex binds to the ICRs via the ANK domain of G9a or WIZ, and maintains the imprinted DNA methylation by recriuting DNMTs via the ANK domain of GLP. (B) In dmANKm ES cells, without forming a dimer with GLP, the stabilities of both the dmANKm G9a and GLP are reduced. However, it is likely that via the interaction with WIZ, the dmANKm G9a monomer still binds to the ICRs. Although there are some GLP/GLP complex exist *in vivo*, the local concentration of the DNMTs is insufficient for the maintenance of imprinted DNA methylation. The black and white dots represent for the methylated and unmethylated cytosines within the ICRs, respectively. The oragne dots represent for the H3K9me2, whereas the green dots represent for the H3K9me0/1.

Furthermore, 5meC MeDIP results also showed that some of the ICRs, such as *Snrpn*, *Peg13* and *GnasXL*, still displayed normal patterns of imprinted DNA methylation in dmANKm ES cells, One possibility could be GLP/GLP homodimers play a role in the maintenance of imprinted DNA methylation at limited imprinted loci, since in theory, GLP/GLP homodimers should also be able to interact and recruit DNMTs to ICRs in dmANKm ES cells. In addition, another possibility could be the existence of trace amount of G9a/GLP heterodimers prefers to bind to the limited number of ICRs and maintain imprinted DNA methylation at these loci, since endogenous wild-type G9a is not removed completely by shRNA in dmANKm ES cells.

Chapter 8 Discussion

8.1 DNA methylation and imprinted DNA methylation

In animals, DNA methylation refers to an addition of a methyl group to the 5-position carbon within cytosine pyrimidine ring. From plants to mammals, the proportion of methylated cytosine within the genome varies dramatically among species. Even for a given species, methylated cytosine is not evenly distributed within the genome. Take mouse genome for instance, repetitive elements, such as the LINEs, SINEs and IAPs, are highly methylated, whereas there is no DNA methylation at promoters of house keeping genes. Since DNA methylation is essential for the normal development and is also associated with a number of crucial biological processes (Smith and Meissner, 2013), genomic regions that contain differentially methylated cytosine have attracted increasing attention.

A differentially methylated region (DMR) refers to a locus within the genome, which contains patterns of DNA methylation that vary between cell types or alleles of the same gene. For instance, some tumour-specific DMRs are only identified in cancer cells, but not in healthy tissues (Doi et al., 2009). In contrast, another category of DMRs, which displays allele-specific differentially methylated patterns, is found in all types of tissues at imprinting control regions (ICRs). Since removal of DNA methylation from these DMRs leads to aberrant expression of adjacent imprinted genes, DNA methylation presents at these DMRs plays an essential *cis*-acting regulatory role in the regulation of imprinted genes (Ziller et al., 2013).

As discussed previously, imprinted DNA methylation is established during the developing of gametes (Figure 1.4) (Davis et al., 1999; Lucifero et al., 2002) and unlike the DNA methylation that located at non-imprinted loci, once established, imprinted

DNA methylation is stably maintained throughout life, even during DNA demethylation that occurs in pre-implantation embryos (Lionel et al., 2010). Since imprinted DNA methylation is essential for accurate expression of imprinted genes and abnormal expression of imprinted genes leads to clinical phenotypes and human disorders, it is essential to investigate the underlying mechanisms that involved in the maintenance of imprinted DNA methylation.

Previous studies have found that DNMTs and other chromatin and/or DNA binding proteins play a role in the maintenance of imprinted DNA methylation (Okano et al., 1999; Chen et al., 2003; Hirasawa et al., 2008; Li et al., 2008; Liang et al., 2002; Nakamura et al., 2007; Nakamura et al., 2012; Quenneville et al., 2011). The experiments described in this thesis identify that G9a/GLP complex of histone methyltransferases as an additional important component in the maintenance of genomic imprinting, specifically in mouse ES cells. In course of my work, I characterised in detail underlying mechanisms by which G9a/GLP complex maintains imprinted DNA methylation in ES cells.

8.2 Impairment of imprinted DNA methylation in either *G9a* or *GLP* deficient ES cells

Experiments carried out in our lab have identified that approximate 170 promoters showed hypomethylated specifically in *G9a-/-* ES cells and among these promoters, some of the maternal imprinted loci were involved (Figure 1.9, Ausma Termanis, unpublished). To confirm this observation, I validated the microarray results and found that DNA methylation is lost from nearly all of the investigated ICRs in *G9a-/-* ES cells compared to wild-type cells (Figure 3.1 and 3.2). In addition, the loss of DNA methylation at ICRs is also apparent when *Glp* is absent from the ES cells (Figure 3.3).

Since G9a/GLP heterodimer serves as the predominant form of G9a complex in ES cells (Ueda et al., 2002), these observations suggest that both H3K9me2 associated histone methyltransferases, G9a and GLP, are required for the protection of imprinted DNA methylation.

Since *G9a-/-* ES cells are derived from wild-type ES cells by gene conversions (Tachibana et al., 2002), it is likely that these ES cells carried normal patterns of imprinted DNA methylation originally. In addition, since *de novo* DNA methylation at germline ICRs occurs during the developing of gametes (Figure 1.4) (Davis et al., 1999; Lucifero et al., 2002), loss of imprinted DNA methylation in *G9a-/-* ES cells also suggested that G9a plays a role in the maintenance of imprinted DNA methylation in ES cells. To confirm this hypothesis, I knocked down *G9a* in wild-type ES cells and found that imprinted DNA methylation is reduced in shG9a ES cells and reduction of imprinted DNA methylation is not due to prolonged culturing of ES cells *in vitro* (Figure 4.5 and 4.6). Furthermore, to ensure that the reduction of imprinted DNA methylation is independent of genetic background of ES cells, I also knocked down *G9a* in another separate ES cell line (Figure 4.7) and found that imprinted DNA methylation is also reduced from all the investigated ICRs (Figure 4.8A and B). These observations indicate that the maintenance of imprinted DNA methylation by G9a is independent of genetic background of ES cells.

A previous study suggested that G9a plays a role in the protection of imprinted DNA methylation at Prader-Willi syndrome-imprinting centre (PWS-IC) in ES cells (Xin et al., 2003). My experiments extend these observations and demonstrate that G9a is required for the maintenance of imprinted DNA methylation at nearly all the ICRs in ES cells. As G9a and GLP serve as the main modifiers for the establishment of H3K9me1/2

in ES cells (Tachibana et al., 2002; Peters et al., 2003; Rice et al., 2003) and STELLA, an H3K9me2 binding protein, has been found to protect DNA methylation that present at the maternally derived genome and at certain paternal ICRs in zygotes (Nakamura et al., 2007; Nakamura et al., 2012), one hypothesis for the maintenance of imprinted DNA methylation by G9a/GLP complex was that G9a and GLP indirectly maintain the imprinted DNA methylation at ICRs and other regions of the genome through establishment of H3K9me2, which attracts STELLA and/or other proteins involved in the heterochromatin compaction that may protect against de-methyltransferase activities, such as TET enzymes. On the other hand, since G9a and GLP are able to directly interact with DNMTs *in vitro* and *in vivo* (Esteve et al., 2006; Epsztejn-Litman et al., 2008) and DNMTs are essential for the maintenance of imprinted DNA methylation (Tucker et al., 1996; Kaneda et al., 2004), another hypothesis for protection of imprinted DNA methylation by G9a/GLP complex was that it occurs through direct recruitment of DNMTs by G9a and/or GLP to ICRs in ES cells.

8.3 Aberrant expression of the imprinted genes in *G9a-/-* ES cells

Previous studies have shown that DNA methylation at ICRs is required for accurate regulation of imprinted genes (Li et al., 1993). Since my observations have shown that imprinted DNA methylation is either lost or reduced in *G9a* knockout and knockdown ES cells, it would be expected that the expression levels of imprinted genes would also change in these *G9a* deficient ES cells.

Most imprinted genes are not expressed in ES cells, but transcription commences upon differentiation. Therefore Ausma Termanis, a former PhD student in our lab investigated the expression of imprinted genes in wild-type and *G9a-/-* ES cells differentiated into embryonic bodies after withdrawal of leukemia inhibitory factor (LIF).

These experiments showed that several imprinted genes from *Dlk1/Dio3* and *Kcnq1* imprinting clusters displayed aberrant expression, *Dlk1*, *Rtl1*, *Rian*, *Dio3*, *Kcnq1*, *Kcnq1ot1*, *Cdkn1c*, *Phlda2* and *Osbpl5* (Ausma Termanis, unpublished). This observation is consistent with previous studies, which indicated that in mouse placenta, the presence of histone methyltransferase G9a is required for silencing of *Kcnq1* cluster on the paternally derived chromosome (Wagschal et al., 2008). Taken together, these observations demonstrate that loss of imprinted DNA methylation in *G9a* deficient ES cells results in aberrant expression of imprinted genes.

8.4 G9a is not essential for the maintenance of the imprinted DNA methylation in MEFs and early embryos

As my work has shown that histone methyltransferases G9a and GLP are essential for the maintenance of imprinted DNA methylation in ES cells, it would be important to investigate whether imprinted DNA methylation in somatic cells and embryos also requires G9a/GLP complex. To investigate whether G9a plays a similar role in mouse embryonic fibroblasts (MEFs), we generated the stable cell line expressing G9a shRNA and found that imprinted DNA methylation was stably maintained at ICRs (Irina Stancheva and Tuo Zhang, unpublished). The possible explanation for this observation is that imprinted DNA methylation might be maintained by GLP/GLP homodimer, since, unlike in shG9a ES cells, the levels of GLP were not dependent on G9a in shG9a MEFs and remained similar between shG9a and shCont MEFs (Tuo Zhang, unpublished). As G9a/GLP heterodimer serves as the predominant form of G9a complex in ES cells (Tachibana et al., 2005; Ueda et al., 2006) and my previous observations have shown that deletion or reduction of G9a also leads to significant degradation of GLP in ES cells (Figure 4.1 and 4.7), the stable levels of GLP in shG9a MEFs raises another question:

what is the predominant form of G9a and/or GLP complexes in MEFs and how is the formation of homo-/heterodimers regulated in different cell types? In other words, do G9a and GLP prefer to form homodimer rather than the heterodimer in MEFs and why? To answer these questions, G9a pulls down combined with Western blotting assays and mass spectrometry would need to be carried out from extracts of both ES cells and MEFs. Furthermore, to investigate whether GLP/GLP homodimers play a role in the maintenance of imprinted DNA methylation in shG9a MEFs, MEFs deficient in both G9a and GLP have to be generated followed by the investigations of imprinted DNA methylation.

8.5 G9a is not essential for the maintenance of the imprinted DNA methylation in early embryos

In mammals, the presence of H3K9 methylations always co-exists with DNA methylation. One of the best examples is germline DMR, since my study and previous studies have demonstrated that both H3K9me2 and H3K9me3 exist at the imprinted loci within the genome (Tuo Zhang, unpublished; Li et al., 2008; Quenneville et al., 2011; Wagschal et al., 2007; Nakamura et al., 2007). As deletion of H3K9 methyltransferases causes reduction of both H3K9me2/3 and imprinted DNA methylation (see chapter 3 and 4), a role for these methyltransferases in coupling the maintenance of DNA methylation with heterochromatin in ES cells has been proposed. By using a G9a inhibitor, I found that the role for G9a in the maintenance of imprinted DNA methylation in ES cells is independent of its catalytic activity (see chapter 5).

However, in *G9a-/-* embryos, my results have shown that imprinted loci still contain normal patterns of DNA methylation (data not shown here) and previous study that carried out by Xin *et al* has shown that removal of G9a from embryos led to

abnormal expression, but not the patterns of DNA methylation at imprinted loci (Xin et al., 2003). In addition, although the levels of H3K9me2 and H3K9me3 were impaired in *G9a* knockout embryos, the placenta-specific imprinted genes still contains normal patterns of DNA methylation (Wagschal et al., 2007). Taken together, these observations suggest that in embryos, G9a is not essential for the maintenance of imprinted DNA methylation or at least, in addition to G9a, some other factors might also play a role in the maintenance of imprinted DNA methylation. Since ES cells are generally used as the models for understanding early development, these observations also raised the question, to what extent ES cells could reflect the early developing embryos and do ES cells serve as the good example for the investigation of genomic imprinting?

Another possible explanation for the stable imprinted DNA methylation in *G9a* knockout embryos is that passive DNA demethylation that requires many cell divisions may have occurred in these embryos. As embryonic cells in E6.5 embryos undergo fewer cell divisions than ES cells under selection in culture, they could maintain imprinted DNA methylation more successfully than ES cells. To investigate how the DNA methylation lost from ICRs and other regions of the genome in *G9a* deficient ES cells, G9a has to be removed from DNA demethylation associated protein deficient ES cells, such as TET1, TET2, AID or TDG, followed by the investigations of imprinted and non-imprinted DNA methylation.

8.6 Neither the wild-type nor the catalytically inactive G9a can restore the imprinted DNA methylation in *G9a-/-* ES cells

To investigate whether loss of imprinted DNA methylation can be restored, I investigated the patterns of imprinted DNA methylation in *G9a-/-* ES cells, in which either wild-type or catalytically inactive G9a were expressed from stably integrated

plasmids (Figure 3.4A) and found that neither wild-type nor catalytically inactive G9a could restore imprinted DNA methylation in *G9a-/-* ES cells (Figure 3.4B, 3.4C and 3.5). This observation is consistent with previous studies (Tucker et al., 1996) and taken together, these observations indicate that once lost, imprinted DNA methylation could not be restored in ES or somatic cells.

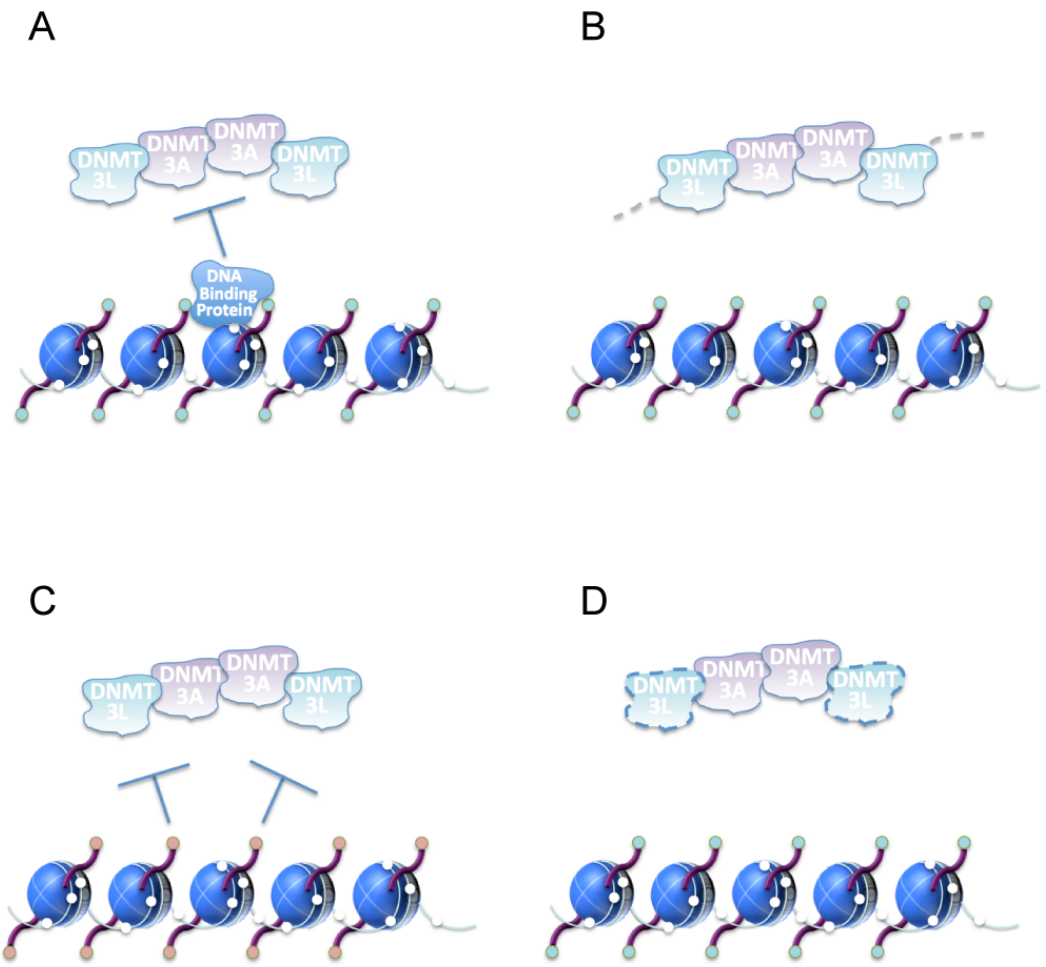


Figure 8.1 Hypothesises of failure to reestablish the imprinted DNA methylation in *G9a-/-* ES cells.

(A) Once the imprinted DNA methylation is lost, the unmethylated DNA binding protein could occupy the methylated allele at the ICRs and prevent the DNMTs to reestablish the DNA methylation. (B) The absence of some specific non-coding transcripts might lead to the failure to reestablish the imprinted DNA methylation in *G9a-/-* ES cells. The dash line represents for the absence of some non-coding transcripts. (C) Gaining of H3K4me2/3 might prevent the DNMTs to bind to the methylated allele at the ICRs and reestablish the imprinted DNA methylation. The orange dots on the histone tail represent for the H3K4me2/3. (D) The absence of the full-length DNMT3L might cause the failure to reestablish the imprinted DNA methylation in *G9a-/-* ES cells. The white dots represent for the ummethylated cytosines, whereas the green dots represent for the histone modifications apart from the H3K4me2/3.

There are several possibilities that may explain why imprinted DNA methylation cannot be restored in *G9a-/-* ES cells (Figure 8.1). First, once imprinted DNA methylation is lost, proteins that bind to the unmethylated DNA may now occupy both ICRs and thus prevent DNMTs or some other factors to restore imprinted DNA methylation. Second, it has been suggested that non-coding transcripts expressed in immature gametes play a role in the establishment of imprinted DNA methylation (Chotalia et al., 2009). Therefore, the failure to re-establish imprinted DNA methylation might be due to absence of such non-coding transcripts in ES cells. Third, since both DNMT3A and DNMT3L are required for the establishment of imprinted DNA methylation in gametes (Arnaud et al., 2006; Kaneda et al., 2004) and full-length DNMT3L seems to exists only in the growing oocytes (Bourc'his et al., 2001), insufficient amount of DNMT3A/DNMT3L complex might also prevent re-establishment of imprinted DNA methylation in *G9a-/-* ES cells. Finally, changes of histone modifications at ICRs might also serve as a barrier for the re-establishment of imprinted DNA methylation, since DNMT3L is unable to bind to chromatin that methylated at H3K4 (Bock et al., 2011). To confirm this hypothesis, H3K4me2 or H3K4me3 ChIP assays have to be carried out at ICRs in *G9a-/-* ES cells.

8.7 DNA methylation is present only on the H3K9me2-marked allele at imprinted loci in ES cells

Previous study has shown that positioning of H3K9 methylation at ICRs is parent-specific and associated with the presence of imprinted DNA methylation (Umlauf et al, 2004). In addition, Quenneville *et al* have shown that H3K9me3 is selectively presents on the methylated allele at ICRs (Quenneville et al., 2011). To investigate whether H3K9me2 coincides with imprinted DNA methylation in ES cells and further ask

whether G9a could discriminate methylated from unmethylated allele, I performed H3K9me2 ChIP combined with bisulfite DNA sequencing and found that in ES cells, DNA methylation is present only on H3K9me2-marked allele at ICRs (Figure 5.1B). This observation indicates that G9a/GLP complex could either somehow discriminate methylated allele from unmethylated one or unmethylated allele is protected from the action of G9a and DNMTs. Furthermore, since my previous observations have shown that imprinted DNA methylation is impaired in *G9a* deficient ES cells, taken together, these observations also suggests that G9a plays a direct role in the maintenance of imprinted DNA methylation in ES cells.

The presence of H3K9me2 on the methylated allele at ICRs raised another question: which is deposited first: imprinted DNA methylation or H3K9me2? Since a zinc finger protein WIZ is another stable component of G9a/GLP complex (Tachibana et al., 2005; Ueda et al., 2006) and it is able to bind to DNA directly (Burak Ozkan, unpublished), it is possible that G9a/GLP complex is selectively recruited to the methylated allele at ICRs via a direct interaction between WIZ and DNA. In this scenario, DNA methylation must be established prior to H3K9me2 and WIZ would function as a factor, which facilitates recruitment of G9a/GLP complex and allow discrimination of methylated alleles from unmethylated ones. To address this hypothesis, DNA binding affinity of WIZ towards methylated and unmethylated DNA must be examined.

On the other hand, since some non-coding transcripts might also play a role in the establishment of imprinted DNA methylation (Kelsey and Feil, 2013) and G9a has been found to interact with such transcripts directly (Nagano et al., 2008; Pandey et al., 2008), it is also possible that during the maturation of gametes, the G9a/GLP complex is selectively recruited by non-coding transcripts to the allele, which has to be methylated

later. Therefore, in this scenario, H3K9me2 would be established prior or concomitantly with imprinted DNA methylation. Whether in addition to the maintenance of DNA methylation, G9a/GLP complex also plays a role in the establishment of DNA methylation at ICRs in ES cells is unknown. To address this question, G9a and/or GLP must be depleted conditionally from developing gametes.

8.8 G9a maintains the imprinted DNA methylation independently of its catalytic activity

My work has shown that imprinted DNA methylation is selectively present on H3K9me2-marked allele at ICRs. As discussed above, since my previous work also showed that imprinted DNA methylation is impaired in *G9a* deficient ES cells and G9a and GLP are the main writers for H3K9me2 (Tachibana et al., 2005; Ueda et al., 2006), taken together, these observations indicate that G9a/GLP complex may play a direct role in the maintenance of imprinted DNA methylation in ES cells. To investigate whether the catalytic activity of G9a is essential for the maintenance of imprinted DNA methylation in ES cells, I employed a G9a inhibitor and found that imprinted DNA methylation is stably maintained in G9a inhibitor treated ES cells (Figure 5.3 and 5.4). This observation suggests that G9a maintains imprinted DNA methylation independently of its catalytic activity. In other words, H3K9me2 is not essential for the maintenance of imprinted DNA methylation by G9a/GLP complex in ES cells.

Furthermore, I also found that G9a is still bound to chromatin when the level of H3K9me2 is dramatically reduced (Figure 5.5). Together with previous studies, which showed that G9a could interact directly with DNMTs (Esteve et al., 2006; Epsztejn-Litman et al., 2008), these observations suggest that G9a protects imprinted DNA methylation by recruitment of DNMTs directly to ICRs via its ANK domain. In addition,

this observation raises a question that how G9a binds to ICRs in G9a inhibitor-treated ES cells. Since the C2H2 zinc finger protein, WIZ, is another stable component of G9a/GLP complex (Tachibana et al., 2005; Ueda et al., 2006) and it is able to bind to DNA directly (Burak Ozkan, unpublished), WIZ might play a role in the recruitment of G9a/GLP complex to ICRs in G9a inhibitor treated ES cells. Whether there is a sequence specific motif within ICRs that WIZ prefers to bind to and whether WIZ prefers to bind to methylated or unmethylated DNA is not yet known. Addressing this question would be important for understanding recruitment of G9a/GLP complex to DNA/chromatin.

In addition, other chromatin or DNA binding proteins might also play a role in the recruitment of G9a/GLP complex to ICRs, since mass spectrometry experiments carried out on G9a that purified from ES cells indicate that, apart from WIZ, other potential chromatin or DNA binding proteins are present in G9a/GLP complex (Burak Ozkan, unpublished). Moreover, long non-coding RNAs (lncRNA) might also play a role in the recruitment of G9a/GLP complex to ICRs, since it has been shown that long non-coding transcripts, such as *Airn* and *Kcnq1ot1*, which are encoded by imprinted genes within two separate imprinting clusters (ICs), could directly interact with G9a protein (Nagano et al., 2008; Pandey et al., 2008).

8.9 Disruption of the interaction between the ANK domain of G9a and H3K9 methylation has no effect on the patterns of imprinted DNA methylation

To investigate whether the recognition and binding of G9a to H3K9me1/2 via its ANK domain is essential for the maintenance of imprinted DNA methylation in ES cells, I generated ANKm ES cell line (Figure 6.1, 6.2 and 6.3) and found that DNA methylation is stable at the majority of ICRs in these cells (Figure 6.5 and 6.6). This observation

indicates that interaction between ANK domain of G9a and H3K9 methylation is not required for the maintenance of imprinted DNA methylation in ES cells. Since ANKm G9a still interacts with DNMTs (Figure 6.7A) and compared with the control cells, a similar amount of DNMTs interact with ANKm G9a (Figure 6.7B), it is likely that without binding to H3K9me1/2 via its ANK domain, ANKm G9a is still able to recruit DNMTs to ICRs and maintain imprinted DNA methylation in ES cells.

Additionally, I also found that the majority of ANKm G9a is tightly bound to chromatin (Figure 6.4). This observation indicates that, apart from recognising and binding to H3K9 methylation via its ANK domain, G9a indirectly binds to chromatin via some other mechanisms. Since ANKm G9a still forms a complex with GLP (Figure 6.7A) and GLP displays similar binding affinity towards methylated H3 peptides to that of G9a (Collins et al., 2008), the ANK domain of GLP might be responsible for indirect recruitment of ANKm G9a to chromatin.

On the other hand, as discussed earlier a C2H2 zinc finger protein WIZ might also play a role in the indirect recruitment of ANKm G9a to chromatin. Further experiments need to be carried out in order to investigate whether disruption of interaction between ANK domain of G9a and H3K9 methylation influences the stability of ANKm-G9a/GLP/WIZ tri-molecular complex in ES cells.

8.10 Removal of the ANK domain from G9a protein reduces the imprinted DNA methylation in ES cells

To investigate whether G9a maintains imprinted DNA methylation via its ANK domain-dependent recruitment of DNMTs to ICRs, I generated ANKdel ES cell line (Figure 6.1, 6.2 and 6.3) and found that DNA methylation is reduced at all the investigated ICRs (Figure 6.5 and 6.6). This observation indicates that deletion of ANK

domain from G9a protein leads to impairment of imprinted DNA methylation in ES cells. Since ANKdel G9a still interacts with GLP (Figure 6.7A) and GLP also contains an ANK domain (Collins et al., 2008), reduction of imprinted DNA methylation in ANKdel ES cells also indicates that G9a/GLP complex, containing of one ANK domain, is not able to maintain normal patterns of imprinted DNA methylation in ES cells.

In ANKdel ES cells, I also found that ANKdel G9a still interacts with DNMTs (Figure 6.7A). Considering that GLP also contains an ANK domain and displays similar chromatin binding patterns to those of G9a (Collins et al., 2008), it is likely that ANKdel G9a indirectly interacts with DNMTs via ANK domain of GLP. In addition, compared with wild-type (shR) and ANKm forms of G9a, I found that less DNMT3A and DNMT1 interacted with ANKdel form of G9a (Figure 6.7B). Therefore, the local concentration of DNMTs, which are recruited to ICRs by G9a/GLP complex, might play an essential role in the maintenance of imprinted DNA methylation in ES cells.

Moreover, I also found that the majority of ANKdel G9a is still tightly bound to chromatin (Figure 6.4). This observation indicates that deletion of ANK domain from G9a has no effect on chromatin binding stability of G9a/GLP complex. As discussed above, it is likely that ANKdel-G9a/GLP complex binds to chromatin via ANK domain of GLP or WIZ.

8.11 Dimerization of G9a and GLP is essential for the maintenance of imprinted DNA methylation in ES cells

To investigate whether ANKm G9a maintains imprinted DNA methylation via the ANK domain of GLP, and further study whether the dimerization between G9a and GLP is essential for the maintenance of imprinted DNA methylation in ES cells, I generated dmANKm ES cell line (Figure 7.1, 7.2 and 7.3) and found that without forming a dimer

with GLP, dmANKm G9a is unable to maintain imprinted DNA methylation. These observations suggest that ANKm G9a maintains imprinted DNA methylation via ANK domain of GLP and dimerization of G9a and GLP is essential for the maintenance of imprinted DNA methylation in ES cells.

In addition, I found that dmANKm G9a still binds to the chromatin (Figure 7.4), which raised the question of how does dmANKm G9a bind to chromatin. Furthermore, I also found that dmANKm G9a lost interaction with DNMTs (Figure 7.5). However, in both dmANKm and dmANKdel ES cells, GLP still interacts with DNMTs (Figure 7.5), although the levels of GLP are significantly reduced in these cell lines (Figure 7.3 A and B). These observations indicate that the reduction of the imprinted DNA methylation in dmANKm ES cells could be caused by insufficient amount of DNMTs that are recruited to ICRs by G9a monomer or GLP/GLP homodimer, although in theory, the remaining GLP/GLP homodimer could interact with DNMTs and recruit them to ICRs. To confirm this hypothesis, ChIP assays of DNMTs at the ICRs in the ANKm and dmANKm ES cells have to be carried out to provide a conclusive answer.

8.12 G9a is also essential for the maintenance of DNA methylation at non-imprinted loci.

H3K9 methylations function as the repressive marks for gene expression and these marks is widespread within the genome. For instances, both H3K9me2 and H3K9me3 are present at nuclear peripheral (Barski et al., 2007) and centromere heterochromatin regions (Lam et al., 2006). In mammals, the main histone methyltransferase for the establishment and maintenance of H3K9me2 is G9a. In mouse ES cells, since G9a protein, but not its catalytic activity, is required for the maintenance of DNA methylation at repetitive elements (Dong et al., 2008) and my results indicate

that removal of G9a also had no effect on the patterns of imprinted DNA methylation (Tuo Zhang, unpublished), these observations suggest that G9a protein *per se* somehow is sufficient for the maintenance of DNA methylation that coexists with H3K9me2. Since several studies have demonstrated that through its N-terminal domain or ANK domain, G9a protein is able to directly interact with DNMTs (Esteve et al., 2006; Epsztejn-Litman et al, 2008), it is likely that G9a maintains H3K9me2-associated DNA methylation via recruitment of DNMTs to its target loci.

Actually, in my study, apart from imprinted loci, I found that some other genomic regions also displayed hypo-methylation in *G9a* deficient ES cells, since compared with wild-type ES cells, the level of global DNA methylation reduced 50% in *G9a-/-* ES cells (data not shown here) and the microarray data indicated that 170 promoter showed hypo-methylation in *G9a-/-* ES cells, among which only 8 were maternal ICRs (Ausma Termanis, unpublished). These observations are consistent with previous studies, which indicates that G9a also plays a role for the maintenance of DNA methylation at repetitive elements within the genome (Dong et al., 2008).

The mechanisms that involved in the maintenance of DNA methylation by G9a at different genomic loci might variant. At nuclear periphery, the role for G9a in the maintenance of DNA methylation seems associated with lamina A, since in human cells, lamina-associated chromatin domains are suppressed and highly methylated with H3K9me2 (Guelen et al., 2008). By contrast, since knockdown of *G9a* causes not only reduction of H3K9me2, but also decreased in HP1 binding at some of the promoters in human cells (EI Gazzar et al., 2008), HP1 might also facilitate G9a in the maintenance of DNA methylation. However, the mechanisms that involved in the maintenance of imprinted DNA methylation by G9a might be different from the above mechanisms,

since after fertilization, imprinted loci are the unique regions that escape from global DNA demethylation. To decipher whether G9a maintains imprinted DNA methylation depending on the presence of lamina A or HP1, Co-IP assays and ChIP assays at imprinted loci that using antibodies against G9a, lamina A or HP1 are required.

8.13 Summary

Taken together, my work revealed that in addition to the establishment and maintenance of H3K9 methylation, the histone methyltransferases G9a and GLP also play an important role in the maintenance of imprinted DNA methylation in ES cells. In addition, I also revealed that the presence of H3K9me2 at ICRs is not essential for the maintenance of genomic imprinting. It is likely that G9a/GLP complex maintains imprinted DNA methylation via the ANK domain-dependent recruitment of DNMTs to ICRs in ES cells.

Previous studies have shown that in ES cells, zinc finger protein ZFP57 is able to bind to the methylated cytosine within TGCCGC motif and that ZFP57 and its partner KAP1 are required for the maintenance of imprinted DNA methylation via recruitment of histone methyltransferase SETDB1, which exclusively establishes the H3K9me3 on the methylated allele at the ICRs (Quenneville et al., 2011). In my study, since both the patterns of imprinted DNA methylation and the levels of H3K9me2/3 at ICRs are impaired in *G9a* deficient ES cells (Chapter 3 and 4), my observations also suggest that binding of ZFP57 at the ICRs would be also dependent on the maintenance of imprinted DNA methylation by histone methyltransferase G9a. To decipher the link between ZFP57 and G9a, the chromatin binding patterns of ZFP57 and SETDB1 have to be analysed in *G9a* deficient ES cells.

Abbreviations

AML	Acute myeloid leukemia
ANK	Ankyrin repeat
ANKdel G9a	ANK domain deleted G9a
ANKm G9a	ANK domain mutant G9a
AR	Androgen receptor
AS	Angelman syndrome
AS-IC	Angelman syndrome imprinting center
BWS	Beckwith-Wiedemann syndrome
Cdkn1c	Cyclin dependent kinase inhibitor 1C
CGI	CpG island
ChIP	Chromatin immunoprecipitation
CLBC	Claudin-low breast cancer
Co-IP	Co-immunoprecipitation
Dlk1	Delta-like homolog 1
dmANKdel G9a	Dimerization-deficient ANKdel G9a
dmANKm G9a	Dimerization-deficient ANKm G9a
DMR	Differentially methylated region
DNMT1	DNA methyltransferase 1
DNMT1o	Oocyte-specific isoform of DNMT1
DNMT3A	DNA methyltransferases 3A
DNMT3B	DNA methyltransferases 3B
DNMT3L	DNA methyltransferases 3L

DNMTs	DNA methyltransferases
E11.5/9.5/6.5	Embryonic day 11.5/9.5/6.5
EMT	Epithelial-mesenchymal transition
Epigenetic mark	Post-translational protein modifications or post-replication nucleotide modifications, which may change expression status of genes without altering DNA sequences
Epigenetic phenomenon	Inheritable changes of phenotype(s) caused by epigenetic marks but not the alternation of DNA sequences
ES cell	Embryonic stem cell
EST	Expressed sequence tag
FBS	Fetal bovine serum
FSHD	Facioscapulohumeral muscular dystrophy
HKMTs	Histone lysine methyltransferases
HP1 gamma	Heterochromatin protein 1 gamma
HPTMs	Histone posttranslational modifications
ICF syndrome	Immunodeficiency, centromere instability and facial anomalies syndrome
ICM	Inner cell mass
Imprinting cluster (IC)	A genomic region which contains imprinted genes that are mono-allelically expressed in the somatic tissue
Imprinting control region (ICR)	A genomic region that contains parent-of-origin-specific epigenetic marks and is required for the regulation of mono-allelic expression of imprinted genes that within the same imprinting cluster
Imprinted genes	Genes whose expression is determined by the parental inherited allele-specific patterns of epigenetic marks

KDMs	Histone lysine demethylases
KMTs	Lysine methyltransferases
KYP	Kryptonite
LIF	Leukemia inhibitory factor
LINEs	Long interspersed repeats
lncRNAs	Long non-coding RNAs
LSD1	Lysine-specific demethylase 1
Maternal ICR	Imprinting control region that inherited from the maternally derived chromosomes
MBD1	Methylated DNA binding protein 1
MeDIP	Methylated DNA immunoprecipitation
MEFs	Mouse embryonic fibroblasts
MEM	Minimum essential medium
MLL	Mixed lineage leukemia
MTD	Methyltransferase domain
NDRs	Nucleosome-depleted regions
NLS	Nuclear localization signal
Paternal ICR	Imprinting control region that inherited from the paternally derived chromosomes
PBD	Proliferating cell nuclear antigen-binding domain
PBS	Phosphate buffered saline
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PGCs	Primordial germ cells
PHD	Plant homeo-domain

PHD	Polybromo homology domain
Pol II	RNA polymerase II
PRC1	Polycomb repressive complex 1
PRC2	Polycomb repressive complex 2
PSG	Penicillin-streptomycin-glutamine
PTHMs	Post-translational histone modifications
PWS	Prader-Willi syndrome
PWS-IC	Prader-Willi syndrome imprinting center
qPCR	Quantitative PCR
RT-qPCR	Reverse transcription quantitative PCR
SAM	*S*-adenosyl-*L*-methionine
shCont	Non-targeting control shRNA
shG9a	shRNA knockdown of *G9a*
shRNA	short hairpin RNA
SINEs	Short interspersed repeats
snoRNA	Small nuclear RNA
SNPs	Single nucleotide polymorphisms
SRS	Silver-Russell syndrome
TDG	Thymine DNA glycosylase
TE	Trophectoderm
TET	Ten-eleven translocation
TF	Transcription factor
TND	Transient neonatal diabetes

tRNA-Asp	Anticodon loop of the aspartic acid transfer RNA
TSS	Transcription start site
UPD	Uniparental disomy
Xi	Inactive X chromosome
Xpr	X-pairing region

Appendix I

Sequences of G9a_ANKm ORF

atgcggggtctgccgagagggaggggctgatgcgggcccgggggcgggggcgtgcg
gcccccacgggcggccgcggccgcggtcgggggggcgcccaccgagggcgaggtagg
ccccgaagcctgctctcgctgcccagggcccaggcgtcttgggcccccagctgcct
gccgggctgaccggcccccggttccttgtctcccctcccaggggaggcccccgct
gagatgggggcgctgctgctggagaaggagccccgaggagccgccgagagagttcat
agctctttgggggacacccctcagagtgaggagacccttcccaaggccaaccccgac
tccttggagcctgccggcccctcctctccggcctctgtcactgtcaccgtcggcgat
gaggggctgacacccctgtcggggccgcatcactcatcggggacgaacccgagagc
ctggagggagatgggggtcgcatcgtgctgggccatgccacaaagtcgttcccctct
tcccccagcaaggggggtgcctgtcccagtcgggccaaaatgtcaatgacaggggca
ggaaagtcgcccccctcggtccagagtttggccatgaggctgttgagcatgcccggg
gcccagggagctgcaactgctgggcctgaaccctctccggcaacaactgccgcccag
gaggggcagcccaaagtgcaccgagcccggaaaaccatgtccaaacctagcaacgga
cagcctccaatccctgagaagcggcccctgaagtccagcatttccgcatgagtgat
gacatgcatctggggaaggtgacttcagatgtggccaaaaggaggaagctgaactct
ggtagcctgtccgaggacttgggctctgccgggggctcaggagatataatcctggag
aagggagagcccaggcccctggaggagtgggagacggtggtgggcgatgacttcagc
ctgtactatgatgcgtactctgtggatgagcgggtggactctgacagcaagtctgaa
gtcgaagctctagctgaacagttgagtgaggaggaggaggaggaagaggaggaagaa
gaagaagaggaggaggaggaggaagaggaggaggaggaagaagaggacgaggagtcg
ggcaatcagtcagacaggagcggttctagtggccggcgcaaggccaagaagaaatgg
cggaaagacagcccgtgggtgaagccatctagaaaacggcggaaacgagagcctccg
agggccaaggagccaagaggagtgaatggtgtgggttcctcagggcccagtgagtac
atggaggttcctctggggtccctggagctgcccagcgaggggaccctctccccaac
cacgctggggtctccaatgacacgtcttcactggagacagaacgcgggtttgaggag
ctgcccctctgcagctgccgcatggaggctcccaagattgaccgcatcagcgagaga
gcagggcacaagtgcatggccacagagagtgtggatggagagctcctgggctgcaat
gctgccatccttaagcgggagaccatgcggccgtctagccgcgtggcgctgatggtg
ctctgtgaggcccatcgagcccgcatggtcaagcaccattgctgcccgggctgcggc
tacttctgcacagcgggcaccttcctggaatgccaccccgactttcgtgtagctcac
cgcttccataaggcctgcgtatcccagctcaatgggatggtcttctgtccccactgt
ggagaggatgcctcagaggcccaggaggtgaccattcctcggggcgatgggggaaca
cccccaattggcaccgcagctcctgctctgccacccctggcacatgatgccccaggg
cgagcggatacctcccagcctagcgcccgaatgcgagggcatggagagccgcggcgc
ccgccctgtgatcccctggctgacaccatcgacagctcagggccttcactgactctg
cctaatgggggctgcctctccgctgtgggtctgcccccagggccgggcagggaagcc
ctggaaaaagccttggtcatccaggagtctgagaggcggaagaagctgcgattccac
ccacggcagctgtacctgtcggtgaagcaggggagctgcagaaggtgatccttatg
ctgttagacaacctggaccccaacttccagagcgaccagcagagcaagcgcacgccc
ctgcacgcggccgcccagaaggggtcggtagagatctgtcatgtgctgctgcaggca
ggagccaacatcaatgccgtagataagcaacaacgcacgccactaatggaggccgtg
gtgaacaaccacctggaggtggcacgctacatggtgcagttaggtggctgtgtctac
agcaaggaagaggatggctccacctgtctacatcatgcagccaaaattgggaacttg
gaaatggtcagcctgctactgagcacaggacaggtggacgtcaatgcccaggacagt
gggggctggacgcccatcatcgctgcagccgctcacaagcacatcgatgtgattcgt

atgctgctgacccggggtgccgatgtcaccctgactgacaatgaggaaaacatctgc
ctgcactgggcctccttcacgggtagtgccgccatcgctgaggtccttctgaatgcc
cagtgtgatctccatgctgtcaactaccatggggacacgcccctgcacatagccgcc
agggagagctaccatgactgtgttctgttgttcctgtctcgtggagccaaccctgag
cttcggaacaaagaaggagacacggcatgggatctgaccccagagcgctctgatgtg
tggtttgcactgcagctcaatcgaaagcttaggcttggggtagggaaccgggctgtc
cgcaccgagaagatcatctgccgggacgtagcccgaggctatgagaatgtacccatc
ccctgtgtcaatggtgtggatggggagccgtgcccggaggactacaagtacatctct
gagaactgcgagacatcgaccatgaacatcgaccgcaacatcacccatctgcagcac
tgcacgtgtgtggatgactgctccagctccaattgcctatgtggtcagctcagtatc
cgatgctggtatgacaaggacgggcggctgctccaggagtttaacaagatcgagccc
cccctgatctttgagtgtaaccaggcatgctcctgctggagaagctgcaagaaccgc
gtggtgcagagcggcatcaaggtacggctgcagctctaccggactgccaagatgggc
tgggggtccgagccttgcagaccatcccccagggcacgttcatctgcgagtatgta
ggagagctgatctctgatgccgaggctgatgtgagagaggatgattcttacctcttc
gatttagataacaaggatggcgaggtttactgcattgatgcccgttactatggcaac
atcagccgattcattaaccacctgtgtgaccccaacatcatccctgtccgggttttc
atgctgcaccaagatctacggttcccacgcattgccttcttcagctccagggacatc
cggactggggaggagctgggctttgactacggtgaccgattctgggacatcaagagc
aagtatttcacctgccagtgtggctctgagaagtgcaagcattcagcggaggccatc
gccctggagcagagccgcctggcccggctggacccccacccggagctgctccctgac
ctcagctccctgcccccccatcaacacctgaggactcttaaaatccaggccgggcact
gcccttcagacatttctccatcagagaccccagtaaggcctggaaggtcgatggccc
ctctcccagagctggtttctcactgggagtgcaagtgacttcagggctggccttccc
cactgagcctggcctcagttagctgattgaagttgggcctctgccagctgattttct
gtgttctcaataaatgttgggtttggtaaaaaaaaa

Appendix II

Sequences of G9a_ANKdel ORF

atgcggggtctgccgagagggaggggctgatgcgggcccgggggcgggggcgtgcg
gcccccacgggcggccgcggccgcggtcgggggggcgcccaccgagggcgaggtagg
ccccgaagcctgctctcgctgcccagggcccaggcgtcttgggcccccagctgcct
gccgggctgaccggcccccggttccttgtctcccctcccaggggaggcccccgct
gagatgggggcgctgctgctggagaaggagccccgaggagccgccgagagagttcat
agctctttgggggacacccctcagagtgaggagacccttcccaaggccaaccccgac
tccttggagcctgccggcccctcctctccggcctctgtcactgtcaccgtcggcgat
gaggggctgacacccctgtcggggccgcatcactcatcggggacgaacccgagagc
ctggagggagatgggggtcgcatcgtgctgggccatgccacaaagtcgttcccctct
tcccccagcaaggggggtgcctgtcccagtcgggccaaaatgtcaatgacaggggca
ggaaagtcgcccccctcggtccagagtttggccatgaggctgttgagcatgcccggg
gcccagggagctgcaactgctgggcctgaaccctctccggcaacaactgccgcccag
gaggggcagcccaaagtgcaccgagcccggaaaaccatgtccaaacctagcaacgga
cagcctccaatccctgagaagcggcccctgaagtccagcatttccgcatgagtgat
gacatgcatctggggaaggtgacttcagatgtggccaaaaggaggaagctgaactct
ggtagcctgtccgaggacttgggctctgccgggggctcaggagatataatcctggag
aagggagagcccaggcccctggaggagtgggagacggtggtgggcgatgacttcagc
ctgtactatgatgcgtactctgtggatgagcgggtggactctgacagcaagtctgaa
gtcgaagctctagctgaacagttgagtgaggaggaggaggaggaagaggaggaagaa
gaagaagaggaggaggaggaggaagaggaggaggaggaagaagaggacgaggagtcg
ggcaatcagtcagacaggagcggttctagtggccggcgcaaggccaagaagaaatgg
cggaaagacagcccgtgggtgaagccatctagaaaacggcggaaacgagagcctccg
agggccaaggagccaagaggagtgaatggtgtgggttcctcagggcccagtgagtac
atggaggttcctctggggtccctggagctgcccagcgaggggaccctctccccaac
cacgctggggtctccaatgacacgtcttcactggagacagaacgcgggtttgaggag
ctgcccctctgcagctgccgcatggaggctcccaagattgaccgcatcagcgagaga
gcagggcacaagtgcatggccacagagagtgtggatggagagctcctgggctgcaat
gctgccatccttaagcgggagaccatgcggccgtctagccgcgtggcgctgatggtg
ctctgtgaggcccatcgagcccgcatggtcaagcaccattgctgcccgggctgcggc
tacttctgcacagcgggcaccttcctggaatgccaccccgactttcgtgtagctcac
cgcttccataaggcctgcgtatcccagctcaatgggatggtcttctgtccccactgt
ggagaggatgcctcagaggcccaggaggtgaccattcctcggggcgatgggggaaca
cccccaattggcaccgcagctcctgctctgccacccctggcacatgatgccccaggg
cgagcggatacctcccagcctagcgcccgaatgcgagggcatggagagccgcggcgc
ccgccctgtgatcccctggctgacaccatcgacagctcagggccttcactgactctg
cctaatgggggctgcctctccgctgtgggtctgcccccagggccgggcagggaagcc
ctggaaaaagccttggtcatcctgaccccagagcgctctgatgtgtggtttgcactg
cagctcaatcgaaagcttaggcttggggtagggaaccgggctgtccgcaccgagaag
atcatctgccgggacgtagcccgaggctatgagaatgtacccatcccctgtgtcaat
ggtgtggatggggagccgtgcccggaggactacaagtacatctctgagaactgcgag
acatcgaccatgaacatcgaccgcaacatcacccatctgcagcactgcacgtgtgtg
gatgactgctccagctccaattgcctatgtggtcagctcagtatccgatgctggtat
gacaaggacgggcggctgctccaggagtttaacaagatcgagcccccctgatcttt
gagtgtaaccaggcatgctcctgctggagaagctgcaagaaccgcgtggtgcagagc
ggcatcaaggtacggctgcagctctaccggactgccaagatgggctgggggtccga

gccttgcagaccatcccccagggcacgttcatctgcgagtatgtaggagagctgatc
tctgatgccgaggctgatgtgagagaggatgattcttacctcttcgatttagataac
aaggatggcgaggtttactgcattgatgcccgttactatggcaacatcagccgattc
attaaccacctgtgtgaccccaacatcatccctgtccgggttttcatgctgcaccaa
gatctacggttcccacgcattgccttcttcagctccagggacatccggactggggag
gagctgggctttgactacggtgaccgattctgggacatcaagagcaagtatttcacc
tgccagtgtggctctgagaagtgcaagcattcagcggaggccatcgccctggagcag
agccgcctggcccggctggacccccacccggagctgctccctgacctcagctccctg
ccccccatcaacacctgaggactcttaaaatccaggccgggcactgcccttcagaca
tttctccatcagagaccccagtaaggcctggaaggtcgatggcccctctcccagagc
tggtttctcactgggagtgcaagtgacttcagggctggccttccccactgagcctgg
cctcagttagctgattgaagttgggcctctgccagctgattttctgtgttctcaata
aatgttgggtttggtaaaaaaaaa

Appendix III

Sequences of G9a_dmANKm ORF

atgcggggtctgccgagagggaggggctgatgcgggcccgggggcgggggcgtgcg
gcccccacgggcggccgcggccgcggtcgggggggcgcccaccgagggcgaggtagg
ccccgaagcctgctctcgctgcccagggcccaggcgtcttgggcccccagctgcct
gccgggctgaccggcccccggttccttgtctcccctcccaggggaggcccccgct
gagatgggggcgctgctgctggagaaggagccccgaggagccgccgagagagttcat
agctctttgggggacacccctcagagtgaggagacccttcccaaggccaaccccgac
tccttggagcctgccggcccctcctctccggcctctgtcactgtcaccgtcggcgat
gaggggctgacacccctgtcggggccgcatcactcatcggggacgaacccgagagc
ctggagggagatgggggtcgcatcgtgctgggccatgccacaaagtcgttcccctct
tcccccagcaaggggggtgcctgtcccagtcgggccaaaatgtcaatgacaggggca
ggaaagtcgcccccctcggtccagagtttggccatgaggctgttgagcatgcccggg
gcccagggagctgcaactgctgggcctgaaccctctccggcaacaactgccgcccag
gaggggcagcccaaagtgcaccgagcccggaaaaccatgtccaaacctagcaacgga
cagcctccaatccctgagaagcggcccctgaagtccagcatttccgcatgagtgat
gacatgcatctggggaaggtgacttcagatgtggccaaaaggaggaagctgaactct
ggtagcctgtccgaggacttgggctctgccgggggctcaggagatataatcctggag
aagggagagcccaggcccctggaggagtgggagacggtggtgggcgatgacttcagc
ctgtactatgatgcgtactctgtggatgagcgggtggactctgacagcaagtctgaa
gtcgaagctctagctgaacagttgagtgaggaggaggaggaggaagaggaggaagaa
gaagaagaggaggaggaggaggaagaggaggaggaggaagaagaggacgaggagtcg
ggcaatcagtcagacaggagcggttctagtggccggcgcaaggccaagaagaaatgg
cggaaagacagcccgtgggtgaagccatctagaaaacggcggaaacgagagcctccg
agggccaaggagccaagaggagtgaatggtgtgggttcctcagggcccagtgagtac
atggaggttcctctggggtccctggagctgcccagcgaggggaccctctccccaac
cacgctggggtctccaatgacacgtcttcactggagacagaacgcgggtttgaggag
ctgcccctctgcagctgccgcatggaggctcccaagattgaccgcatcagcgagaga
gcagggcacaagtgcatggccacagagagtgtggatggagagctcctgggctgcaat
gctgccatccttaagcgggagaccatgcggccgtctagccgcgtggcgctgatggtg
ctctgtgaggcccatcgagcccgcatggtcaagcaccattgctgcccgggctgcggc
tacttctgcacagcgggcaccttcctggaatgccaccccgactttcgtgtagctcac
cgcttccataaggcctgcgtatcccagctcaatgggatggtcttctgtccccactgt
ggagaggatgcctcagaggcccaggaggtgaccattcctcggggcgatgggggaaca
cccccaattggcaccgcagctcctgctctgccacccctggcacatgatgccccaggg
cgagcggatacctcccagcctagcgcccgaatgcgagggcatggagagccgcggcgc
ccgccctgtgatcccctggctgacaccatcgacagctcagggccttcactgactctg
cctaatgggggctgcctctccgctgtgggtctgcccccagggccgggcagggaagcc
ctggaaaaagccttggtcatccaggagtctgagaggcggaagaagctgcgattccac
ccacggcagctgtacctgtcggtgaagcaggggagctgcagaaggtgatccttatg
ctgttagacaacctggaccccaacttccagagcgaccagcagagcaagcgcacgccc
ctgcacgcggccgcccagaaggggtcggtagagatctgtcatgtgctgctgcaggca
ggagccaacatcaatgccgtagataagcaacaacgcacgccactaatggaggccgtg
gtgaacaaccacctggaggtggcacgctacatggtgcagttaggtggctgtgtctac
agcaaggaagaggatggctccacctgtctacatcatgcagccaaaattgggaacttg
gaaatggtcagcctgctactgagcacaggacaggtggacgtcaatgcccaggacagt
gggggctggacgcccatcatcgctgcagccgctcacaagcacatcgatgtgattcgt

atgctgctgacccggggtgccgatgtcaccctgactgacaatgaggaaaacatctgc
ctgcactgggcctccttcacgggtagtgccgccatcgctgaggtccttctgaatgcc
cagtgtgatctccatgctgtcaactaccatggggacacgcccctgcacatagccgcc
agggagagctaccatgactgtgttctgttgttcctgtctcgtggagccaaccctgag
cttcggaacaaagaaggagacacggcatgggatctgaccccagagcgctctgatgtg
tggtttgcactgcagctcaatcgaaagcttaggcttggggtagggaaccgggctgtc
cgcaccgagaagatcatctgccgggacgtagcccgaggctatgagaatgtacccatc
ccctgtgtcaatggtgtggatggggagccgtgcccggaggactacaagtacatctct
gagaactgcgagacatcgaccatgaacatcgaccgcaacatcacccatctgcagcac
tgcacgtgtgtggatgactgctccagctccaattgcctatgtggtcagctcagtatc
cgatgctggtatgacaaggacgggcggctgctccaggagtttaacaagatcgagccc
cccctgatctttgagtgtaaccaggcatgctcctgctggagaagctgcaagaaccgc
gtggtgcagagcggcatcaaggtacggctgcagctctaccggactgccaagatgggc
tgggggtccgagccttgcagaccatcccccagggcacgttcatctgcgagtatgta
ggagagctgatctctgatgccgaggctgatgtgagagaggatgattcttacctcttc
gatttagataacaaggatggcgaggtttactgcattgatgcccgttactatggcaac
atcagccgattcattctcgagctgtgtgaccccaacatcatccctgtccgggttttc
atgctgcaccaagatctacggttcccacgcattgccttcttcagctccagggacatc
cggactggggaggagctgggctttgactacggtgaccgattctgggacatcaagagc
aagtatttcacctgccagtgtggctctgagaagtgcaagcattcagcggaggccatc
gccctggagcagagccgcctggcccggctggacccccacccggagctgctccctgac
ctcagctccctgcccccccatcaacacctgaggactcttaaaatccaggccgggcact
gcccttcagacatttctccatcagagaccccagtaaggcctggaaggtcgatggccc
ctctcccagagctggtttctcactgggagtgcaagtgacttcagggctggccttccc
cactgagcctggcctcagttagctgattgaagttgggcctctgccagctgattttct
gtgttctcaataaatgttgggtttggtaaaaaaaaa

Appendix IV

Sequences of G9a_dmANKdel ORF

atgcggggtctgccgagagggaggggctgatgcgggcccgggggcgggggcgtgcg
gcccccacgggcggccgcggccgcggtcgggggggcgcccaccgagggcgaggtagg
ccccgaagcctgctctcgctgcccagggcccaggcgtcttgggcccccagctgcct
gccgggctgaccggcccccggttccttgtctcccctcccaggggaggcccccgct
gagatgggggcgctgctgctggagaaggagccccgaggagccgccgagagagttcat
agctctttgggggacacccctcagagtgaggagacccttcccaaggccaaccccgac
tccttggagcctgccggcccctcctctccggcctctgtcactgtcaccgtcggcgat
gaggggctgacacccctgtcggggccgcatcactcatcggggacgaacccgagagc
ctggagggagatgggggtcgcatcgtgctgggccatgccacaaagtcgttcccctct
tcccccagcaaggggggtgcctgtcccagtcgggccaaaatgtcaatgacaggggca
ggaaagtcgcccccctcggtccagagtttggccatgaggctgttgagcatgcccggg
gcccagggagctgcaactgctgggcctgaaccctctccggcaacaactgccgcccag
gaggggcagcccaaagtgcaccgagcccggaaaaccatgtccaaacctagcaacgga
cagcctccaatccctgagaagcggccccctgaagtccagcatttccgcatgagtgat
gacatgcatctggggaaggtgacttcagatgtggccaaaaggaggaagctgaactct
ggtagcctgtccgaggacttgggctctgccgggggctcaggagatataatcctggag
aagggagagcccaggcccctggaggagtgggagacggtggtgggcgatgacttcagc
ctgtactatgatgcgtactctgtggatgagcgggtggactctgacagcaagtctgaa
gtcgaagctctagctgaacagttgagtgaggaggaggaggaggaagaggaggaagaa
gaagaagaggaggaggaggaggaagaggaggaggaggaagaagaggacgaggagtcg
ggcaatcagtcagacaggagcggttctagtggccggcgcaaggccaagaagaaatgg
cggaaagacagcccgtgggtgaagccatctagaaaacggcggaaacgagagcctccg
agggccaaggagccaagaggagtgaatggtgtgggttcctcagggcccagtgagtac
atggaggttcctctggggtccctggagctgcccagcgaggggaccctctccccaac
cacgctggggtctccaatgacacgtcttcactggagacagaacgcgggtttgaggag
ctgcccctctgcagctgccgcatggaggctcccaagattgaccgcatcagcgagaga
gcagggcacaagtgcatggccacagagagtgtggatggagagctcctgggctgcaat
gctgccatccttaagcgggagaccatgcggccgtctagccgcgtggcgctgatggtg
ctctgtgaggcccatcgagcccgcatggtcaagcaccattgctgcccgggctgcggc
tacttctgcacagcgggcaccttcctggaatgccaccccgactttcgtgtagctcac
cgcttccataaggcctgcgtatcccagctcaatgggatggtcttctgtccccactgt
ggagaggatgcctcagaggcccaggaggtgaccattcctcggggcgatgggggaaca
cccccaattggcaccgcagctcctgctctgccacccctggcacatgatgccccaggg
cgagcggatacctcccagcctagcgcccgaatgcgagggcatggagagccgcggcgc
ccgccctgtgatcccctggctgacaccatcgacagctcagggccttcactgactctg
cctaatgggggctgcctctccgctgtgggtctgcccccagggccgggcagggaagcc
ctggaaaaagccttggtcatcctgaccccagagcgctctgatgtgtggtttgcactg
cagctcaatcgaaagcttaggcttggggtagggaaccgggctgtccgcaccgagaag
atcatctgccgggacgtagcccgaggctatgagaatgtacccatcccctgtgtcaat
ggtgtggatggggagccgtgcccggaggactacaagtacatctctgagaactgcgag
acatcgaccatgaacatcgaccgcaacatcacccatctgcagcactgcacgtgtgtg
gatgactgctccagctccaattgcctatgtggtcagctcagtatccgatgctggtat
gacaaggacgggcggctgctccaggagtttaacaagatcgagcccccctgatcttt
gagtgtaaccaggcatgctcctgctggagaagctgcaagaaccgcgtggtgcagagc
ggcatcaaggtacggctgcagctctaccggactgccaagatgggctgggggtccga

gccttgcagaccatcccccagggcacgttcatctgcgagtatgtaggagagctgatc
tctgatgccgaggctgatgtgagagaggatgattcttacctcttcgatttagataac
aaggatggcgaggtttactgcattgatgcccgttactatggcaacatcagccgattc
attctcgagctgtgtgaccccaacatcatccctgtccgggttttcatgctgcaccaa
gatctacggttcccacgcattgccttcttcagctccagggacatccggactggggag
gagctgggctttgactacggtgaccgattctgggacatcaagagcaagtatttcacc
tgccagtgtggctctgagaagtgcaagcattcagcggaggccatcgccctggagcag
agccgcctggcccggctggacccccacccggagctgctccctgacctcagctccctg
ccccccatcaacacctgaggactcttaaaatccaggccgggcactgcccttcagaca
tttctccatcagagaccccagtaaggcctggaaggtcgatggcccctctcccagagc
tggtttctcactgggagtgcaagtgacttcagggctggccttccccactgagcctgg
cctcagttagctgattgaagttgggcctctgccagctgattttctgtgttctcaata
aatgttgggtttggtaaaaaaaaa

References

Abu-Amero, S., Monk, D., Frost, J., Preece, M., Stanier, P., and Moore, G.E. (2008). The genetic aetiology of Silver-Russell syndrome. Journal of medical genetics *45*, 193-199.

Ager, E.I., Pask, A.J., Gehring, H.M., Shaw, G., and Renfree, M.B. (2008). Evolution of the CDKN1C-KCNQ1 imprinted domain. BMC evolutionary biology *8*, 163.

Altun, G., Loring, J.F., and Laurent, L.C. (2010). DNA methylation in embryonic stem cells. Journal of cellular biochemistry *109*, 1-6.

Ariel, M., Robinson, E., McCarrey, J.R., and Cedar, H. (1995). Gamete-specific methylation correlates with imprinting of the murine Xist gene. Nature genetics *9*, 312-315.

Arnaud, P., Hata, K., Kaneda, M., Li, E., Sasaki, H., Feil, R., and Kelsey, G. (2006). Stochastic imprinting in the progeny of Dnmt3L-/- females. Human molecular genetics *15*, 589-598.

Athanasiadou, R., de Sousa, D., Myant, K., Merusi, C., Stancheva, I., and Bird, A. (2010). Targeting of de novo DNA methylation throughout the Oct-4 gene regulatory region in differentiating embryonic stem cells. PloS one *5*, e9937.

Avery, O.T., Macleod, C.M., and McCarty, M. (1944). Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. The Journal of experimental medicine *79*, 137-158.

Babak, T., Deveale, B., Armour, C., Raymond, C., Cleary, M.A., van der Kooy, D., Johnson, J.M., and Lim, L.P. (2008). Global survey of genomic imprinting by transcriptome sequencing. Current biology : CB *18*, 1735-1741.

Bacolla, A., Pradhan, S., Roberts, R.J., and Wells, R.D. (1999). Recombinant human DNA (cytosine-5) methyltransferase. II. Steady-state kinetics reveal allosteric activation by methylated dna. The Journal of biological chemistry *274*, 33011-33019.

Bannister, A.J., Schneider, R., and Kouzarides, T. (2002). Histone methylation: dynamic or static? Cell *109*, 801-806.

Barlow, D.P., Stoger, R., Herrmann, B.G., Saito, K., and Schweifer, N. (1991). The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. Nature *349*, 84-87.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.

Barth, T.K., and Imhof, A. (2010). Fast signals and slow marks: the dynamics of histone modifications. Trends in biochemical sciences *35*, 618-626.

Bartolomei, M.S., Zemel, S., and Tilghman, S.M. (1991). Parental imprinting of the mouse H19 gene. Nature *351*, 153-155.

Bauer, U.M., Daujat, S., Nielsen, S.J., Nightingale, K., and Kouzarides, T. (2002). Methylation at arginine 17 of histone H3 is linked to gene activation. EMBO reports *3*, 39-44.

Beard, W.A., and Wilson, S.H. (1995). Purification and domain-mapping of mammalian DNA polymerase beta. Methods in enzymology *262*, 98-107.

Beisson, J., and Sonneborn, T.M. (1965). Cytoplasmic Inheritance of the Organization of the Cell Cortex in Paramecium Aurelia. Proceedings of the National Academy of Sciences of the United States of America *53*, 275-282.

Belkina, A.C., and Denis, G.V. (2012). BET domain co-regulators in obesity, inflammation and cancer. Nature reviews. Cancer *12*, 465-477.

Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature *405*, 482-485.

Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., 3rd, Gingeras, T.R., *et al.* (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. Cell *120*, 169-181.

Bernstein, E., Duncan, E.M., Masui, O., Gil, J., Heard, E., and Allis, C.D. (2006). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. Molecular and cellular biology *26*, 2560-2569.

Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. Journal of molecular biology *203*, 971-983.

Bestor, T.H. (1988). Cloning of a mammalian DNA methyltransferase. Gene *74*, 9-12.

Bickmore, W.A., and van Steensel, B. (2013). Genome architecture: domain organization of interphase chromosomes. Cell *152*, 1270-1284.

Bielinska, B., Blaydes, S.M., Buiting, K., Yang, T., Krajewska-Walasek, M., Horsthemke, B., and Brannan, C.I. (2000). De novo deletions of SNRPN exon 1 in early human and mouse embryos result in a paternal to maternal imprint switch. Nature genetics *25*, 74-78.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes & development *16*, 6-21.

Bird, A., Taggart, M., Frommer, M., Miller, O.J., and Macleod, D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell *40*, 91-99.

Bird, A.P. (1986). CpG-rich islands and the function of DNA methylation. Nature *321*, 209-213.

Bock, I., Dhayalan, A., Kudithipudi, S., Brandt, O., Rathert, P., and Jeltsch, A. (2011). Detailed specificity analysis of antibodies binding to modified histone tails with peptide arrays. Epigenetics : official journal of the DNA Methylation Society *6*, 256-263.

Bourc'his, D., and Bestor, T.H. (2006). Origins of extreme sexual dimorphism in genomic imprinting. Cytogenetic and genome research *113*, 36-40.

Bourc'his, D., Le Bourhis, D., Patin, D., Niveleau, A., Comizzoli, P., Renard, J.P., and Viegas-Pequignot, E. (2001). Delayed and incomplete reprogramming of chromosome methylation patterns in bovine cloned embryos. Current biology : CB *11*, 1542-1546.

Braunstein, M., Rose, A.B., Holmes, S.G., Allis, C.D., and Broach, J.R. (1993). Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. Genes & development *7*, 592-604.

Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell *84*, 843-851.

Buehr, M. (1997). The primordial germ cells of mammals: some current perspectives. Experimental cell research *232*, 194-207.

Buiting, K. (2010). Prader-Willi syndrome and Angelman syndrome. American journal of medical genetics. Part C, Seminars in medical genetics *154C*, 365-376.

Cang, S., Feng, J., Konno, S., Han, L., Liu, K., Sharma, S.C., Choudhury, M., and Chiao, J.W. (2009). Deficient histone acetylation and excessive deacetylase activity as epigenomic marks of prostate cancer cells. International journal of oncology *35*, 1417-1422.

Cankovic, M., Mikkelsen, T., Rosenblum, M.L., and Zarbo, R.J. (2007). A simplified laboratory validated assay for MGMT promoter hypermethylation analysis of glioma specimens from formalin-fixed paraffin-embedded tissue. Laboratory investigation; a journal of technical methods and pathology *87*, 392-397.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-1043.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P.*, et al.* (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581-592.

Cassidy, S.B., and Driscoll, D.J. (2009). Prader-Willi syndrome. European journal of human genetics : EJHG *17*, 3-13.

Chang, Y., Zhang, X., Horton, J.R., Upadhyay, A.K., Spannhoff, A., Liu, J., Snyder, J.P., Bedford, M.T., and Cheng, X. (2009). Structural basis for G9a-like protein lysine methyltransferase inhibition by BIX-01294. Nature structural & molecular biology *16*, 312-317.

Changizi, V., Takavar, A., Babakhani, A., and Sohrabi, M. (2008). Scatter correction for heart SPECT images using TEW method. Journal of applied clinical medical physics / American College of Medical Physics *9*, 2767.

Chedin, F., Lieber, M.R., and Hsieh, C.L. (2002). The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. Proceedings of the National Academy of Sciences of the United States of America *99*, 16916-16921.

Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., and Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Molecular and cellular biology *23*, 5594-5605.

Cheng, X. (1995). DNA modification by methyltransferases. Current opinion in structural biology *5*, 4-10.

Cheng, X., and Blumenthal, R.M. (2008). Mammalian DNA methyltransferases: a structural perspective. Structure *16*, 341-350.

Chess, A. (2013). Random and non-random monoallelic expression. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology *38*, 55-61.

Cheung, P., Tanner, K.G., Cheung, W.L., Sassone-Corsi, P., Denu, J.M., and Allis, C.D. (2000). Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. Molecular cell *5*, 905-915.

Chotalia, M., Smallwood, S.A., Ruf, N., Dawson, C., Lucifero, D., Frontera, M., James, K., Dean, W., and Kelsey, G. (2009). Transcription is required for establishment of germline methylation marks at imprinted genes. Genes & development *23*, 105-117.

Choufani, S., Shuman, C., and Weksberg, R. (2010). Beckwith-Wiedemann syndrome. American journal of medical genetics. Part C, Seminars in medical genetics *154C*, 343-354.

Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M., and Jacobsen, S.E. (2008). Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature *452*, 215-219.

Collins, K.A., Camahort, R., Seidel, C., Gerton, J.L., and Biggins, S. (2007). The overexpression of a Saccharomyces cerevisiae centromeric histone H3 variant mutant protein leads to a defect in kinetochore biorientation. Genetics *175*, 513-525.

Collins, R.E., Northrop, J.P., Horton, J.R., Lee, D.Y., Zhang, X., Stallcup, M.R., and Cheng, X. (2008). The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. Nature structural & molecular biology *15*, 245-250.

da Rocha, S.T., Charalambous, M., Lin, S.P., Gutteridge, I., Ito, Y., Gray, D., Dean, W., and Ferguson-Smith, A.C. (2009). Gene dosage effects of the imprinted delta-like homologue 1 (dlk1/pref1) in development: implications for the evolution of imprinting. PLoS genetics *5*, e1000392.

Danam, R.P., Howell, S.R., Remack, J.S., and Brent, T.P. (2001). Heterogeneous methylation of the O(6)-methylguanine-DNA methyltransferase promoter in immortalized IMR90 cell lines. International journal of oncology *18*, 1187-1193.

Davies, W., Lynn, P.M., Relkovic, D., and Wilkinson, L.S. (2008). Imprinted genes and neuroendocrine function. Frontiers in neuroendocrinology *29*, 413-427.

Davis, T.L., Trasler, J.M., Moss, S.B., Yang, G.J., and Bartolomei, M.S. (1999). Acquisition of the H19 methylation imprint occurs differentially on the parental alleles during spermatogenesis. Genomics *58*, 18-28.

Dean, W., Bowden, L., Aitchison, A., Klose, J., Moore, T., Meneses, J.J., Reik, W., and Feil, R. (1998). Altered imprinted gene methylation and expression in completely ES cell-derived mouse fetuses: association with aberrant phenotypes. Development *125*, 2273-2282.

DeChiara, T.M., Robertson, E.J., and Efstratiadis, A. (1991). Parental imprinting of the mouse insulin-like growth factor II gene. Cell *64*, 849-859.

DeManno, D.A., Cottom, J.E., Kline, M.P., Peters, C.A., Maizels, E.T., and Hunzicker-Dunn, M. (1999). Follicle-stimulating hormone promotes histone H3 phosphorylation on serine-10. Molecular endocrinology *13*, 91-105.

Ding, F., Li, H.H., Zhang, S., Solomon, N.M., Camper, S.A., Cohen, P., and Francke, U. (2008). SnoRNA Snord116 (Pwcr1/MBII-85) deletion causes growth deficiency and hyperphagia in mice. PloS one *3*, e1709.

Dodge, J.E., Kang, Y.K., Beppu, H., Lei, H., and Li, E. (2004). Histone H3-K9 methyltransferase ESET is essential for early development. Molecular and cellular biology *24*, 2478-2486.

Dodge, J.E., Okano, M., Dick, F., Tsujimoto, N., Chen, T., Wang, S., Ueda, Y., Dyson, N., and Li, E. (2005). Inactivation of Dnmt3b in mouse embryonic fibroblasts results in DNA hypomethylation, chromosomal instability, and spontaneous immortalization. The Journal of biological chemistry *280*, 17986-17991.

Doi, A., Park, I.H., Wen, B., Murakami, P., Aryee, M.J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S*., et al.* (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nature genetics *41*, 1350-1353.

Dong, C., Wu, Y., Yao, J., Wang, Y., Yu, Y., Rychahou, P.G., Evers, B.M., and Zhou, B.P. (2012). G9a interacts with Snail and is critical for Snail-mediated E-cadherin repression in human breast cancer. The Journal of clinical investigation *122*, 1469-1486.

Dong, K.B., Maksakova, I.A., Mohn, F., Leung, D., Appanah, R., Lee, S., Yang, H.W., Lam, L.L., Mager, D.L., Schubeler, D*., et al.* (2008). DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. The EMBO journal *27*, 2691-2701.

Easwaran, H.P., Schermelleh, L., Leonhardt, H., and Cardoso, M.C. (2004). Replication-independent chromatin loading of Dnmt1 during G2 and M phases. EMBO reports *5*, 1181-1186.

Eden, A., Gaudet, F., Waghmare, A., and Jaenisch, R. (2003). Chromosomal instability and tumors promoted by DNA hypomethylation. Science *300*, 455.

Ehrlich, M. (2003). Expression of various genes is controlled by DNA methylation during mammalian development. Journal of cellular biochemistry *88*, 899-910.

Ehrlich, M., Gama-Sosa, M.A., Huang, L.H., Midgett, R.M., Kuo, K.C., McCune, R.A., and Gehrke, C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic acids research *10*, 2709-2721.

El Gazzar, M., Yoza, B.K., Chen, X., Hu, J., Hawkins, G.A., and McCall, C.E. (2008). G9a and HP1 couple histone and DNA methylation to TNFalpha transcription silencing during endotoxin tolerance. The Journal of biological chemistry *283*, 32198-32208.

Epsztejn-Litman, S., Feldman, N., Abu-Remaileh, M., Shufaro, Y., Gerson, A., Ueda, J., Deplus, R., Fuks, F., Shinkai, Y., Cedar, H.*, et al.* (2008). De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. Nature structural & molecular biology *15*, 1176-1183.

Espada, J., Ballestar, E., Fraga, M.F., Villar-Garea, A., Juarranz, A., Stockert, J.C., Robertson, K.D., Fuks, F., and Esteller, M. (2004). Human DNA methyltransferase 1 is required for maintenance of the histone H3 modification pattern. The Journal of biological chemistry *279*, 37175-37184.

Esteve, P.O., Chin, H.G., Smallwood, A., Feehery, G.R., Gangisetty, O., Karpf, A.R., Carey, M.F., and Pradhan, S. (2006). Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. Genes & development *20*, 3089-3103.

Fabbrizio, E., El Messaoudi, S., Polanowska, J., Paul, C., Cook, J.R., Lee, J.H., Negre, V., Rousset, M., Pestka, S., Le Cam, A.*, et al.* (2002). Negative regulation of transcription by the type II arginine methyltransferase PRMT5. EMBO reports *3*, 641-645.

Farthing, C.R., Ficz, G., Ng, R.K., Chan, C.F., Andrews, S., Dean, W., Hemberger, M., and Reik, W. (2008). Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. PLoS genetics *4*, e1000116.

Fatemi, M., Hermann, A., Gowher, H., and Jeltsch, A. (2002). Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. European journal of biochemistry / FEBS *269*, 4981-4984.

Fatemi, M., Hermann, A., Pradhan, S., and Jeltsch, A. (2001). The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. Journal of molecular biology *309*, 1189-1199.

Fedoriw, A.M., Stein, P., Svoboda, P., Schultz, R.M., and Bartolomei, M.S. (2004). Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. Science *303*, 238-240.

Feldman, N., Gerson, A., Fang, J., Li, E., Zhang, Y., Shinkai, Y., Cedar, H., and Bergman, Y. (2006). G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nature cell biology *8*, 188-194.

Feng, S.Y., Ota, K., and Ito, T. (2010). A yeast one-hybrid system to screen for methylated DNA-binding proteins. Nucleic acids research *38*, e189.

Filion, G.J., and van Steensel, B. (2010). Reassessing the abundance of H3K9me2 chromatin domains in embryonic stem cells. Nature genetics *42*, 4; author reply 5-6.

Fingerman, I.M., Li, H.C., and Briggs, S.D. (2007). A charge-based interaction between histone H4 and Dot1 is required for H3K79 methylation and telomere silencing: identification of a new trans-histone pathway. Genes & development *21*, 2018-2029.

Fischle, W., Tseng, B.S., Dormann, H.L., Ueberheide, B.M., Garcia, B.A., Shabanowitz, J., Hunt, D.F., Funabiki, H., and Allis, C.D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. Nature *438*, 1116-1122.

Fraga, M.F., Herranz, M., Espada, J., Ballestar, E., Paz, M.F., Ropero, S., Erkek, E., Bozdogan, O., Peinado, H., Niveleau, A.*, et al.* (2004). A mouse skin multistage carcinogenesis model reflects the aberrant DNA methylation patterns of human tumors. Cancer research *64*, 5527-5534.

Frederiks, F., Tzouros, M., Oudgenoeg, G., van Welsem, T., Fornerod, M., Krijgsveld, J., and van Leeuwen, F. (2008). Nonprocessive methylation by Dot1 leads to functional redundancy of histone H3K79 methylation states. Nature structural & molecular biology *15*, 550-557.

Fritsch, L., Robin, P., Mathieu, J.R., Souidi, M., Hinaux, H., Rougeulle, C., Harel-Bellan, A., Ameyar-Zazoua, M., and Ait-Si-Ali, S. (2010). A subset of the histone H3 lysine 9 methyltransferases Suv39h1, G9a, GLP, and SETDB1 participate in a multimeric complex. Molecular cell *37*, 46-56.

Gardner, R.G., Nelson, Z.W., and Gottschling, D.E. (2005). Ubp10/Dot4p regulates the persistence of ubiquitinated histone H2B: distinct roles in telomeric silencing and general chromatin. Molecular and cellular biology *25*, 6123-6139.

Gholivand, M.B., Babakhanian, A., and Rafiee, E. (2008). Determination of Sn(II) and Sn(IV) after mixed micelle-mediated cloud point extraction using alpha-polyoxometalate as a complexing agent by flame atomic absorption spectrometry. Talanta *76*, 503-508.

Gicquel, C., Rossignol, S., Cabrol, S., Houang, M., Steunou, V., Barbu, V., Danton, F., Thibaud, N., Le Merrer, M., Burglen, L.*, et al.* (2005). Epimutation of the telomeric imprinting center region on chromosome 11p15 in Silver-Russell syndrome. Nature genetics *37*, 1003-1007.

Glozak, M.A., Sengupta, N., Zhang, X., and Seto, E. (2005). Acetylation and deacetylation of non-histone proteins. Gene *363*, 15-23.

Goryachkina, K., Babak, S., Burbello, A., Wettemark, B., and Bergman, U. (2008a). Quality use of medicines: a new method of combining antibiotic consumption and sensitivity data--application in a Russian hospital. Pharmacoepidemiology and drug safety *17*, 636-644.

Goryachkina, K., Burbello, A., Boldueva, S., Babak, S., Bergman, U., and Bertilsson, L. (2008b). CYP2D6 is a major determinant of metoprolol disposition and effects in hospitalized Russian patients treated for acute myocardial infarction. European journal of clinical pharmacology *64*, 1163-1173.

Goryachkina, K., Burbello, A., Boldueva, S., Babak, S., Bergman, U., and Bertilsson, L. (2008c). Inhibition of metoprolol metabolism and potentiation of its effects by paroxetine in routinely treated patients with acute myocardial infarction (AMI). European journal of clinical pharmacology *64*, 275-282.

Gowher, H., and Jeltsch, A. (2001). Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites. Journal of molecular biology *309*, 1201-1208.

Gowher, H., Liebert, K., Hermann, A., Xu, G., and Jeltsch, A. (2005a). Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. The Journal of biological chemistry *280*, 13341-13348.

Gowher, H., Stockdale, C.J., Goyal, R., Ferreira, H., Owen-Hughes, T., and Jeltsch, A. (2005b). De novo methylation of nucleosomal DNA by the mammalian Dnmt1 and Dnmt3A DNA methyltransferases. Biochemistry *44*, 9899-9904.

Gruenbaum, Y., Cedar, H., and Razin, A. (1982). Substrate and sequence specificity of a eukaryotic DNA methylase. Nature *295*, 620-622.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W*., et al.* (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature *453*, 948-951.

Guillemot, F., Caspary, T., Tilghman, S.M., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Anderson, D.J., Joyner, A.L., Rossant, J., and Nagy, A. (1995). Genomic imprinting of Mash2, a mouse gene required for trophoblast development. Nature genetics *9*, 235-242.

Guzman, E., Babakhani, A., and Maker, V.K. (2008). Improving outcomes on the ABS Certifying Examination: can monthly mock orals do it? Journal of surgical education *65*, 441-444.

Hahn, M.A., Wu, X., Li, A.X., Hahn, T., and Pfeifer, G.P. (2011). Relationship between gene body DNA methylation and intragenic H3K9me3 and H3K36me3 chromatin marks. PloS one *6*, e18844.

Hata, K., Okano, M., Lei, H., and Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development *129*, 1983-1993.

Hebbes, T.R., Clayton, A.L., Thorne, A.W., and Crane-Robinson, C. (1994). Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. The EMBO journal *13*, 1823-1830.

Hellman, A., and Chess, A. (2007). Gene body-specific methylation on the active X chromosome. Science *315*, 1141-1143.

Hermann, A., Schmitt, S., and Jeltsch, A. (2003). The human Dnmt2 has residual DNA-(cytosine-C5) methyltransferase activity. The Journal of biological chemistry *278*, 31717-31721.

Hershey, A.D., and Chase, M. (1952). Independent functions of viral protein and nucleic acid in growth of bacteriophage. The Journal of general physiology *36*, 39-56.

Hirasawa, R., Chiba, H., Kaneda, M., Tajima, S., Li, E., Jaenisch, R., and Sasaki, H. (2008). Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. Genes & development *22*, 1607-1616.

Holliday, R., and Pugh, J.E. (1975). DNA modification mechanisms and gene activity during development. Science *187*, 226-232.

Horsthemke, B., and Wagstaff, J. (2008). Mechanisms of imprinting of the Prader-Willi/Angelman region. American journal of medical genetics. Part A *146A*, 2041-2052.

Huang, J., Dorsey, J., Chuikov, S., Perez-Burgos, L., Zhang, X., Jenuwein, T., Reinberg, D., and Berger, S.L. (2010). G9a and Glp methylate lysine 373 in the tumor suppressor p53. The Journal of biological chemistry *285*, 9636-9641.

Huang, S., Litt, M., and Felsenfeld, G. (2005). Methylation of histone H4 by arginine methyltransferase PRMT1 is essential in vivo for many subsequent histone modifications. Genes & development *19*, 1885-1893.

Iida, T., Suetake, I., Tajima, S., Morioka, H., Ohta, S., Obuse, C., and Tsurimoto, T. (2002). PCNA clamp facilitates action of DNA cytosine methyltransferase 1 on hemimethylated DNA. Genes to cells : devoted to molecular & cellular mechanisms *7*, 997-1007.

Illingworth, R.S., Gruenewald-Schneider, U., Webb, S., Kerr, A.R., James, K.D., Turner, D.J., Smith, C., Harrison, D.J., Andrews, R., and Bird, A.P. (2010). Orphan CpG islands identify numerous conserved promoters in the mammalian genome. PLoS genetics *6*, e1001134.

Imai, K., Togami, H., and Okamoto, T. (2010). Involvement of histone H3 lysine 9 (H3K9) methyltransferase G9a in the maintenance of HIV-1 latency and its reactivation by BIX01294. The Journal of biological chemistry *285*, 16538-16545.

Isagawa, T., Nagae, G., Shiraki, N., Fujita, T., Sato, N., Ishikawa, S., Kume, S., and Aburatani, H. (2011). DNA methylation profiling of embryonic stem cell differentiation into the three germ layers. PloS one *6*, e26052.

Ishida, M., and Moore, G.E. (2013). The role of imprinted genes in humans. Molecular aspects of medicine *34*, 826-840.

Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature *466*, 1129-1133.

Jabs, E.W., Wolf, S.F., and Migeon, B.R. (1984). Characterization of reiterated human DNA with respect to mammalian X chromosome homology. Somatic cell and molecular genetics *10*, 93-103.

Jeltsch, A. (2002). Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. Chembiochem : a European journal of chemical biology *3*, 274-293.

Jiang, L., Jobst, P., Lai, L., Samuel, M., Ayares, D., Prather, R.S., and Tian, X.C. (2007). Expression levels of growth-regulating imprinted genes in cloned piglets. Cloning and stem cells *9*, 97-106.

Johnson, L.M., Bostick, M., Zhang, X., Kraft, E., Henderson, I., Callis, J., and Jacobsen, S.E. (2007). The SRA methyl-cytosine-binding domain links DNA and histone methylation. Current biology : CB *17*, 379-384.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature reviews. Genetics *13*, 484-492.

Joshi, A.A., and Struhl, K. (2005). Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. Molecular cell *20*, 971-978.

Jurkowska, M., and Bal, J. (2001). [Biomedical science in the era of complete sequence of human genome]. Medycyna wieku rozwojowego *5*, 197-212.

Jurkowski, T.P., Meusburger, M., Phalke, S., Helm, M., Nellen, W., Reuter, G., and Jeltsch, A. (2008). Human DNMT2 methylates tRNA(Asp) molecules using a DNA methyltransferase-like catalytic mechanism. RNA *14*, 1663-1670.

Kacem, S., and Feil, R. (2009). Chromatin mechanisms in genomic imprinting. Mammalian genome : official journal of the International Mammalian Genome Society *20*, 544-556.

Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M., and Saitou, M. (2013). Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. The EMBO journal *32*, 340-353.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004a). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature *429*, 900-903.

Kaneda, M., Sado, T., Hata, K., Okano, M., Tsujimoto, N., Li, E., and Sasaki, H. (2004b). Role of de novo DNA methyltransferases in initiation of genomic imprinting and X-chromosome inactivation. Cold Spring Harbor symposia on quantitative biology *69*, 125-129.

Kannenberg, K., Urban, C., and Binder, G. (2012). Increased incidence of aberrant DNA methylation within diverse imprinted gene loci outside of IGF2/H19 in Silver-Russell syndrome. Clinical genetics *81*, 366-377.

Karachentsev, D., Sarma, K., Reinberg, D., and Steward, R. (2005). PR-Set7-dependent methylation of histone H4 Lys 20 functions in repression of gene expression and is essential for mitosis. Genes & development *19*, 431-435.

Kawahara, M., Wu, Q., Ferguson-Smith, A.C., and Kono, T. (2007). Appropriate expression of imprinted genes on mouse chromosome 12 extends development of bi-maternal embryos to term. FEBS letters *581*, 5178-5184.

Kelly, T.K., Miranda, T.B., Liang, G., Berman, B.P., Lin, J.C., Tanay, A., and Jones, P.A. (2010). H2A.Z maintenance during mitosis reveals nucleosome shifting on mitotically silenced genes. Molecular cell *39*, 901-911.

Kelsey, G., and Feil, R. (2013). New insights into establishment and maintenance of DNA methylation imprints in mammals. Philosophical transactions of the Royal Society of London. Series B, Biological sciences *368*, 20110336.

Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J.*, et al.* (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell *123*, 593-605.

Khasanov, S.A., Babakhanov, G.K., Khasanov, M.S., and Novikov, S.A. (2008). [Improved diagnostic and surgical techniques in adenoid treatment in children]. Vestnik otorinolaringologii, 55-56.

Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S.S., Janssen, H., Amendola, M., Nolen, L.D., Bickmore, W.A., and van Steensel, B. (2013). Single-cell dynamics of genome-nuclear lamina interactions. Cell *153*, 178-192.

Kirmizis, A., Bartley, S.M., Kuzmichev, A., Margueron, R., Reinberg, D., Green, R., and Farnham, P.J. (2004). Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. Genes & development *18*, 1592-1605.

Kizer, K.O., Phatnani, H.P., Shibata, Y., Hall, H., Greenleaf, A.L., and Strahl, B.D. (2005). A novel domain in Set2 mediates RNA polymerase II interaction and

couples histone H3 K36 methylation with transcript elongation. Molecular and cellular biology *25*, 3305-3316.

Klose, R.J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., Tempst, P., Wong, J., and Zhang, Y. (2006). The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. Nature *442*, 312-316.

Kloth, M., Goering, W., Ribarska, T., Arsov, C., Sorensen, K.D., and Schulz, W.A. (2012). The SNP rs6441224 influences transcriptional activity and prognostically relevant hypermethylation of RARRES1 in prostate cancer. International journal of cancer. Journal international du cancer *131*, E897-904.

Kobayashi, H., Sakurai, T., Sato, S., Nakabayashi, K., Hata, K., and Kono, T. (2012). Imprinted DNA methylation reprogramming during early mouse embryogenesis at the Gpr1-Zdbf2 locus is linked to long cis-intergenic transcription. FEBS letters *586*, 827-833.

Kohli, R.M., and Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. Nature *502*, 472-479.

Kohlmaier, A., Savarese, F., Lachner, M., Martens, J., Jenuwein, T., and Wutz, A. (2004). A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. PLoS biology *2*, E171.

Kolpakov, V.V., Bespalova, T.V., Bragin, A.V., Babakin, E.A., Lebedeva, K.A., and Semenov, V.V. (2008). [The concept of typological variability of physiological individuality: I. Intrapopulation diversity of human habitual physical activity and its typological estimation]. Fiziologiia cheloveka *34*, 121-132.

Kondo, Y., Shen, L., Suzuki, S., Kurokawa, T., Masuko, K., Tanaka, Y., Kato, H., Mizuno, Y., Yokoe, M., Sugauchi, F.*, et al.* (2007). Alterations of DNA methylation and histone modifications contribute to gene silencing in hepatocellular carcinomas. Hepatology research : the official journal of the Japan Society of Hepatology *37*, 974-983.

Kou, Y.C., Shao, L., Peng, H.H., Rosetta, R., del Gaudio, D., Wagner, A.F., Al-Hussaini, T.K., and Van den Veyver, I.B. (2008). A recurrent intragenic genomic duplication, other novel mutations in NLRP7 and imprinting defects in recurrent biparental hydatidiform moles. Molecular human reproduction *14*, 33-40.

Kubicek, S., O'Sullivan, R.J., August, E.M., Hickey, E.R., Zhang, Q., Teodoro, M.L., Rea, S., Mechtler, K., Kowalski, J.A., Homon, C.A.*, et al.* (2007). Reversal of H3K9me2 by a small-molecule inhibitor for the G9a histone methyltransferase. Molecular cell *25*, 473-481.

Kurdistani, S.K., and Grunstein, M. (2003). Histone acetylation and deacetylation in yeast. Nature reviews. Molecular cell biology *4*, 276-284.

Lam, A.L., Boivin, C.D., Bonney, C.F., Rudd, M.K., and Sullivan, B.A. (2006). Human centromeric chromatin is a dynamic chromosomal domain that can spread over noncentromeric DNA. Proceedings of the National Academy of Sciences of the United States of America *103*, 4186-4191.

Larsen, S., Rygaard, K., Asnaes, S., and Spang-Thomsen, M. (1992). Northern and Southern blot analysis of human RNA and DNA in autopsy material. APMIS : acta pathologica, microbiologica, et immunologica Scandinavica *100*, 498-502.

Lawson, K.A., Teteak, C.J., Gao, J., Li, N., Hacquebord, J., Ghatan, A., Zielinska-Kwiatkowska, A., Song, G., Chansky, H.A., and Yang, L. (2013a). ESET histone methyltransferase regulates osteoblastic differentiation of mesenchymal stem cells during postnatal bone development. FEBS letters *587*, 3961-3967.

Lawson, K.A., Teteak, C.J., Zou, J., Hacquebord, J., Ghatan, A., Zielinska-Kwiatkowska, A., Fernandes, R.J., Chansky, H.A., and Yang, L. (2013b). Mesenchyme-specific knockout of ESET histone methyltransferase causes ectopic hypertrophy and terminal differentiation of articular chondrocytes. The Journal of biological chemistry *288*, 32119-32125.

Lefebvre, L., Viville, S., Barton, S.C., Ishino, F., Keverne, E.B., and Surani, M.A. (1998). Abnormal maternal behaviour and growth retardation associated with loss of the imprinted gene Mest. Nature genetics *20*, 163-169.

Lehnertz, B., Ueda, Y., Derijck, A.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Current biology : CB *13*, 1192-1200.

Leonhardt, H., Page, A.W., Weier, H.U., and Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell *71*, 865-873.

Leung, D.C., Dong, K.B., Maksakova, I.A., Goyal, P., Appanah, R., Lee, S., Tachibana, M., Shinkai, Y., Lehnertz, B., Mager, D.L.*, et al.* (2011). Lysine methyltransferase G9a is required for de novo DNA methylation and the establishment, but not the maintenance, of proviral silencing. Proceedings of the National Academy of Sciences of the United States of America *108*, 5718-5723.

Li, E., Beard, C., Forster, A.C., Bestor, T.H., and Jaenisch, R. (1993a). DNA methylation, genomic imprinting, and mammalian development. Cold Spring Harbor symposia on quantitative biology *58*, 297-305.

Li, E., Beard, C., and Jaenisch, R. (1993b). Role for DNA methylation in genomic imprinting. Nature *366*, 362-365.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-926.

Li, K.C., Hua, K.T., Lin, Y.S., Su, C.Y., Ko, J.Y., Hsiao, M., Kuo, M.L., and Tan, C.T. (2014). Inhibition of G9a induces DUSP4-dependent autophagic cell death in head and neck squamous cell carcinoma. Molecular cancer *13*, 172.

Li, T., Vu, T.H., Ulaner, G.A., Littman, E., Ling, J.Q., Chen, H.L., Hu, J.F., Behr, B., Giudice, L., and Hoffman, A.R. (2005). IVF results in de novo DNA methylation and histone methylation at an Igf2-H19 imprinting epigenetic switch. Molecular human reproduction *11*, 631-640.

Li, X., Ito, M., Zhou, F., Youngson, N., Zuo, X., Leder, P., and Ferguson-Smith, A.C. (2008). A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. Developmental cell *15*, 547-557.

Liang, G., Chan, M.F., Tomigahara, Y., Tsai, Y.C., Gonzales, F.A., Li, E., Laird, P.W., and Jones, P.A. (2002). Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. Molecular and cellular biology *22*, 480-491.

Liang, G., Lin, J.C., Wei, V., Yoo, C., Cheng, J.C., Nguyen, C.T., Weisenberger, D.J., Egger, G., Takai, D., Gonzales, F.A.*, et al.* (2004). Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. Proceedings of the National Academy of Sciences of the United States of America *101*, 7357-7362.

Lienert, F., Mohn, F., Tiwari, V.K., Baubec, T., Roloff, T.C., Gaidatzis, D., Stadler, M.B., and Schubeler, D. (2011). Genomic prevalence of heterochromatic H3K9me2 and transcription do not discriminate pluripotent from terminally differentiated cells. PLoS genetics *7*, e1002090.

Link, P.A., Gangisetty, O., James, S.R., Woloszynska-Read, A., Tachibana, M., Shinkai, Y., and Karpf, A.R. (2009). Distinct roles for histone methyltransferases G9a and GLP in cancer germ-line antigen gene regulation in human cancer cells and murine embryonic stem cells. Molecular cancer research : MCR *7*, 851-862.

Lister, R., and Ecker, J.R. (2009). Finding the fifth base: genome-wide sequencing of cytosine methylation. Genome research *19*, 959-966.

Litt, M.D., Simpson, M., Gaszner, M., Allis, C.D., and Felsenfeld, G. (2001). Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. Science *293*, 2453-2455.

Liu, Z., Yadav, P.K., Xu, X., Su, J., Chen, C., Tang, M., Lin, H., Yu, J., Qian, J., Yang, P.C.*, et al.* (2011). The increased expression of IL-23 in inflammatory bowel disease promotes intraepithelial and lamina propria lymphocyte inflammatory responses and cytotoxicity. Journal of leukocyte biology *89*, 597-606.

Loh, Y.H., Zhang, W., Chen, X., George, J., and Ng, H.H. (2007). Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells. Genes & development *21*, 2545-2557.

Loyola, A., Bonaldi, T., Roche, D., Imhof, A., and Almouzni, G. (2006). PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. Molecular cell *24*, 309-316.

Lucifero, D., Mertineit, C., Clarke, H.J., Bestor, T.H., and Trasler, J.M. (2002). Methylation dynamics of imprinted genes in mouse germ cells. Genomics *79*, 530-538.

Mackay, D.J., Callaway, J.L., Marks, S.M., White, H.E., Acerini, C.L., Boonen, S.E., Dayanikli, P., Firth, H.V., Goodship, J.A., Haemers, A.P.*, et al.* (2008). Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57. Nature genetics *40*, 949-951.

Marcon, E., Babak, T., Chua, G., Hughes, T., and Moens, P.B. (2008). miRNA and piRNA localization in the male mammalian meiotic nucleus. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *16*, 243-260.

Margot, J.B., Cardoso, M.C., and Leonhardt, H. (2001). Mammalian DNA methyltransferases show different subnuclear distributions. Journal of cellular biochemistry *83*, 373-379.

Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nature reviews. Molecular cell biology *6*, 838-849.

Maze, I., Covington, H.E., 3rd, Dietz, D.M., LaPlant, Q., Renthal, W., Russo, S.J., Mechanic, M., Mouzon, E., Neve, R.L., Haggarty, S.J.*, et al.* (2010). Essential role of the histone methyltransferase G9a in cocaine-induced plasticity. Science *327*, 213-216.

McGarvey, K.M., Fahrner, J.A., Greene, E., Martens, J., Jenuwein, T., and Baylin, S.B. (2006). Silenced tumor suppressor genes reactivated by DNA demethylation do not return to a fully euchromatic chromatin state. Cancer research *66*, 3541-3549.

Metzger, E., Wissmann, M., Yin, N., Muller, J.M., Schneider, R., Peters, A.H., Gunther, T., Buettner, R., and Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature *437*, 436-439.

Montero, L.M., Filipski, J., Gil, P., Capel, J., Martinez-Zapater, J.M., and Salinas, J. (1992). The distribution of 5-methylcytosine in the nuclear genome of plants. Nucleic acids research *20*, 3207-3210.

Moser, D., Ekawardhani, S., Kumsta, R., Palmason, H., Bock, C., Athanassiadou, Z., Lesch, K.P., and Meyer, J. (2009). Functional analysis of a potassium-chloride co-transporter 3 (SLC12A6) promoter polymorphism leading to an additional DNA

methylation site. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology *34*, 458-467.

Muller, H.J., and Altenburg, E. (1930). The Frequency of Translocations Produced by X-Rays in Drosophila. Genetics *15*, 283-311.

Murdoch, S., Seoud, M., Kircheisen, R., Mazhar, B., and Slim, R. (2006). Detailed gene and allele content analysis of three homozygous KIR haplotypes. Tissue antigens *68*, 72-77.

Murray, K. (1964). The Occurrence of Epsilon-N-Methyl Lysine in Histones. Biochemistry *3*, 10-15.

Myant, K., Termanis, A., Sundaram, A.Y., Boe, T., Li, C., Merusi, C., Burrage, J., de Las Heras, J.I., and Stancheva, I. (2011). LSH and G9a/GLP complex are required for developmentally programmed DNA methylation. Genome research *21*, 83-94.

Nagano, T., Mitchell, J.A., Sanz, L.A., Pauler, F.M., Ferguson-Smith, A.C., Feil, R., and Fraser, P. (2008). The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science *322*, 1717-1720.

Nakamura, T., Arai, Y., Umehara, H., Masuhara, M., Kimura, T., Taniguchi, H., Sekimoto, T., Ikawa, M., Yoneda, Y., Okabe, M.*, et al.* (2007). PGC7/Stella protects against DNA demethylation in early embryogenesis. Nature cell biology *9*, 64-71.

Nakamura, T., Liu, Y.J., Nakashima, H., Umehara, H., Inoue, K., Matoba, S., Tachibana, M., Ogura, A., Shinkai, Y., and Nakano, T. (2012). PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. Nature *486*, 415-419.

Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D., and Grewal, S.I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science *292*, 110-113.

Nativio, R., Sparago, A., Ito, Y., Weksberg, R., Riccio, A., and Murrell, A. (2011). Disruption of genomic neighbourhood at the imprinted IGF2-H19 locus in Beckwith-Wiedemann syndrome and Silver-Russell syndrome. Human molecular genetics *20*, 1363-1374.

Nielsen, S.J., Schneider, R., Bauer, U.M., Bannister, A.J., Morrison, A., O'Carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R.E.*, et al.* (2001). Rb targets histone H3 methylation and HP1 to promoters. Nature *412*, 561-565.

Nimura, K., Ishida, C., Koriyama, H., Hata, K., Yamanaka, S., Li, E., Ura, K., and Kaneda, Y. (2006). Dnmt3a2 targets endogenous Dnmt3L to ES cell chromatin and induces regional DNA methylation. Genes to cells : devoted to molecular & cellular mechanisms *11*, 1225-1237.

Noma, K., Allis, C.D., and Grewal, S.I. (2001). Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. Science *293*, 1150-1155.

Oakes, C.C., La Salle, S., Smiraglia, D.J., Robaire, B., and Trasler, J.M. (2007). A unique configuration of genome-wide DNA methylation patterns in the testis. Proceedings of the National Academy of Sciences of the United States of America *104*, 228-233.

Oda, M., Yamagiwa, A., Yamamoto, S., Nakayama, T., Tsumura, A., Sasaki, H., Nakao, K., Li, E., and Okano, M. (2006). DNA methylation regulates long-range gene silencing of an X-linked homeobox gene cluster in a lineage-specific manner. Genes & development *20*, 3382-3394.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247-257.

Okano, M., Xie, S., and Li, E. (1998a). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nature genetics *19*, 219-220.

Okano, M., Xie, S., and Li, E. (1998b). Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. Nucleic acids research *26*, 2536-2540.

Ono, R., Nakamura, K., Inoue, K., Naruse, M., Usami, T., Wakisaka-Saito, N., Hino, T., Suzuki-Migishima, R., Ogonuki, N., Miki, H.*, et al.* (2006). Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. Nature genetics *38*, 101-106.

Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D.*, et al.* (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature *448*, 714-717.

Palmer, L.E., Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V.S., Nascimento, L.U., Dike, S., de la Bastide, M., Martienssen, R.A., and McCombie, W.R. (2003). Maize genome sequencing by methylation filtration. Science *302*, 2115-2117.

Pandey, R.R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L., Komorowski, J., Nagano, T., Mancini-Dinardo, D., and Kanduri, C. (2008). Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Molecular cell *32*, 232-246.

Pedersen, M.T., and Helin, K. (2010). Histone demethylases in development and disease. Trends in cell biology *20*, 662-671.

Peters, A.H., Kubicek, S., Mechtler, K., O'Sullivan, R.J., Derijck, A.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y.*, et al.* (2003). Partitioning

and plasticity of repressive histone methylation states in mammalian chromatin. Molecular cell *12*, 1577-1589.

Petersen-Mahrt, S.K., and Neuberger, M.S. (2003). In vitro deamination of cytosine to uracil in single-stranded DNA by apolipoprotein B editing complex catalytic subunit 1 (APOBEC1). The Journal of biological chemistry *278*, 19583-19586.

Pinheiro, I., Margueron, R., Shukeir, N., Eisold, M., Fritzsch, C., Richter, F.M., Mittler, G., Genoud, C., Goyama, S., Kurokawa, M.*, et al.* (2012). Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity. Cell *150*, 948-960.

Pollard, K.S., Serre, D., Wang, X., Tao, H., Grundberg, E., Hudson, T.J., Clark, A.G., and Frazer, K. (2008). A genome-wide approach to identifying novel-imprinted genes. Human genetics *122*, 625-634.

Pontvianne, F., Blevins, T., and Pikaard, C.S. (2010). Arabidopsis Histone Lysine Methyltransferases. Advances in botanical research *53*, 1-22.

Pradhan, S., Bacolla, A., Wells, R.D., and Roberts, R.J. (1999). Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. The Journal of biological chemistry *274*, 33002-33010.

Qiu, Y., Zhao, Y., Becker, M., John, S., Parekh, B.S., Huang, S., Hendarwanto, A., Martinez, E.D., Chen, Y., Lu, H.*, et al.* (2006). HDAC1 acetylation is linked to progressive modulation of steroid receptor-induced gene transcription. Molecular cell *22*, 669-679.

Quenneville, S., Verde, G., Corsinotti, A., Kapopoulou, A., Jakobsson, J., Offner, S., Baglivo, I., Pedone, P.V., Grimaldi, G., Riccio, A.*, et al.* (2011). In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. Molecular cell *44*, 361-372.

Raghunandan, R., Ruiz-Hidalgo, M., Jia, Y., Ettinger, R., Rudikoff, E., Riggins, P., Farnsworth, R., Tesfaye, A., Laborda, J., and Bauer, S.R. (2008). Dlk1 influences differentiation and function of B lymphocytes. Stem cells and development *17*, 495-507.

Rakyan, V.K., Preis, J., Morgan, H.D., and Whitelaw, E. (2001). The marks, mechanisms and memory of epigenetic states in mammals. The Biochemical journal *356*, 1-10.

Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proceedings of the National Academy of Sciences of the United States of America *97*, 5237-5242.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., *et al.* (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593-599.

Reyes-Dominguez, Y., Narendja, F., Berger, H., Gallmetzer, A., Fernandez-Martin, R., Garcia, I., Scazzocchio, C., and Strauss, J. (2008). Nucleosome positioning and histone H3 acetylation are independent processes in the Aspergillus nidulans prnD-prnB bidirectional promoter. Eukaryotic cell *7*, 656-663.

Rice, J.C., Briggs, S.D., Ueberheide, B., Barber, C.M., Shabanowitz, J., Hunt, D.F., Shinkai, Y., and Allis, C.D. (2003). Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. Molecular cell *12*, 1591-1598.

Riggs, A.D. (1975). X inactivation, differentiation, and DNA methylation. Cytogenetics and cell genetics *14*, 9-25.

Robertson, K.D., Keyomarsi, K., Gonzales, F.A., Velicescu, M., and Jones, P.A. (2000). Differential mRNA expression of the human DNA methyltransferases (DNMTs) 1, 3a and 3b during the G(0)/G(1) to S phase transition in normal and tumor cells. Nucleic acids research *28*, 2108-2113.

Rottach, A., Frauer, C., Pichler, G., Bonapace, I.M., Spada, F., and Leonhardt, H. (2010). The multi-domain protein Np95 connects DNA methylation and histone modification. Nucleic acids research *38*, 1796-1804.

Rountree, M.R., and Selker, E.U. (2010). DNA methylation and the formation of heterochromatin in Neurospora crassa. Heredity *105*, 38-44.

SanMiguel, P., Tikhonov, A., Jin, Y.K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P.S., Edwards, K.J., Lee, M., Avramova, Z., *et al.* (1996). Nested retrotransposons in the intergenic regions of the maize genome. Science *274*, 765-768.

Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proceedings of the National Academy of Sciences of the United States of America *103*, 1412-1417.

Schaefer, A., Sampath, S.C., Intrator, A., Min, A., Gertler, T.S., Surmeier, D.J., Tarakhovsky, A., and Greengard, P. (2009). Control of cognition and adaptive behavior by the GLP/G9a epigenetic suppressor complex. Neuron *64*, 678-691.

Schlegelova, J., Babak, V., Holasova, M., and Dendis, M. (2008). The biofilm-positive Staphylococcus epidermidis isolates in raw materials, foodstuffs and on contact surfaces in processing plants. Folia microbiologica *53*, 500-504.

Schurter, B.T., Koh, S.S., Chen, D., Bunick, G.J., Harp, J.M., Hanson, B.L., Henschen-Edman, A., Mackay, D.R., Stallcup, M.R., and Aswad, D.W. (2001). Methylation of

histone H3 by coactivator-associated arginine methyltransferase 1. Biochemistry *40*, 5747-5756.

Sekigawa, I., Okada, M., Ogasawara, H., Kaneko, H., Hishikawa, T., and Hashimoto, H. (2003). DNA methylation in systemic lupus erythematosus. Lupus *12*, 79-85.

Sekita, Y., Wagatsuma, H., Nakamura, K., Ono, R., Kagami, M., Wakisaka, N., Hino, T., Suzuki-Migishima, R., Kohda, T., Ogura, A*., et al.* (2008). Role of retrotransposon-derived imprinted gene, Rtl1, in the feto-maternal interface of mouse placenta. Nature genetics *40*, 243-248.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K*., et al.* (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature *450*, 908-912.

Shemer, R., Birger, Y., Riggs, A.D., and Razin, A. (1997). Structure of the imprinted mouse Snrpn gene and establishment of its parental-specific methylation pattern. Proceedings of the National Academy of Sciences of the United States of America *94*, 10267-10272.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941-953.

Sibley, C.P., Coan, P.M., Ferguson-Smith, A.C., Dean, W., Hughes, J., Smith, P., Reik, W., Burton, G.J., Fowden, A.L., and Constancia, M. (2004). Placental-specific insulin-like growth factor 2 (Igf2) regulates the diffusional exchange characteristics of the mouse placenta. Proceedings of the National Academy of Sciences of the United States of America *101*, 8204-8208.

Sims, J.K., Houston, S.I., Magazinnik, T., and Rice, J.C. (2006). A trans-tail histone code defined by monomethylated H4 Lys-20 and H3 Lys-9 demarcates distinct regions of silent chromatin. The Journal of biological chemistry *281*, 12760-12766.

Sims, R.J., 3rd, and Reinberg, D. (2006). Histone H3 Lys 4 methylation: caught in a bind? Genes & development *20*, 2779-2786.

Sleutels, F., Zwart, R., and Barlow, D.P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature *415*, 810-813.

Smith, C.R., Mutti, N.S., Jasper, W.C., Naidu, A., Smith, C.D., and Gadau, J. (2012). Patterns of DNA methylation in development, division of labor and hybridization in an ant with genetic caste determination. PloS one *7*, e42433.

Smith, F.M., Holt, L.J., Garfield, A.S., Charalambous, M., Koumanov, F., Perry, M., Bazzani, R., Sheardown, S.A., Hegarty, B.D., Lyons, R.J*., et al.* (2007). Mice with a

disruption of the imprinted Grb10 gene exhibit altered body composition, glucose homeostasis, and insulin signaling during postnatal life. Molecular and cellular biology *27*, 5871-5886.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nature reviews. Genetics *14*, 204-220.

Solter, M., and Sekso, M. (1983). Effect of fasting on posthyperglycemic glucose homeostasis in obesity--experimental model for reactive hypoglycemia. Experimental and clinical endocrinology *81*, 33-40.

Sturtevant, A.H., and Morgan, T.H. (1923). Reverse Mutation of the Bar Gene Correlated with Crossing Over. Science *57*, 746-747.

Surani, M.A., and Barton, S.C. (1983). Development of gynogenetic eggs in the mouse: implications for parthenogenetic embryos. Science *222*, 1034-1036.

Suzuki, K., Kawamura, M., Mineo, M., Shinohara, T., Kataharada, K., Okada, M., Takada, K., Miyawaki, S., and Ohsuzu, F. (2002). The effects of affinity-purified anti-DNA antibodies from patients with systemic lupus erythematosus on the fluorescent antinuclear antibody assay using HEp-2 cells. Clinical chemistry and laboratory medicine : CCLM / FESCC *40*, 46-51.

Szyf, M. (2001). The role of DNA methyltransferase 1 in growth control. Frontiers in bioscience : a journal and virtual library *6*, D599-609.

Tachibana, M., Matsumura, Y., Fukuda, M., Kimura, H., and Shinkai, Y. (2008). G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. The EMBO journal *27*, 2681-2690.

Tachibana, M., Sugimoto, K., Fukushima, T., and Shinkai, Y. (2001). Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. The Journal of biological chemistry *276*, 25309-25317.

Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H.*, et al.* (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes & development *16*, 1779-1791.

Tachibana, M., Ueda, J., Fukuda, M., Takeda, N., Ohta, T., Iwanari, H., Sakihama, T., Kodama, T., Hamakubo, T., and Shinkai, Y. (2005). Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. Genes & development *19*, 815-826.

Takahashi, K., Kobayashi, T., and Kanayama, N. (2000). p57(Kip2) regulates the proper development of labyrinthine and spongiotrophoblasts. Molecular human reproduction *6*, 1019-1025.

Tamaru, H., and Selker, E.U. (2001). A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. Nature *414*, 277-283.

Tamaru, H., and Selker, E.U. (2003). Synthesis of signals for de novo DNA methylation in Neurospora crassa. Molecular and cellular biology *23*, 2379-2394.

Tang, L.Y., Reddy, M.N., Rasheva, V., Lee, T.L., Lin, M.J., Hung, M.S., and Shen, C.K. (2003). The eukaryotic DNMT2 genes encode a new class of cytosine-5 DNA methyltransferases. The Journal of biological chemistry *278*, 33613-33616.

Thorvaldsen, J.L., Duran, K.L., and Bartolomei, M.S. (1998). Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. Genes & development *12*, 3693-3702.

Thuault, S., Gangloff, Y.G., Kirchner, J., Sanders, S., Werten, S., Romier, C., Weil, P.A., and Davidson, I. (2002). Functional analysis of the TFIID-specific yeast TAF4 (yTAF(II)48) reveals an unexpected organization of its histone-fold domain. The Journal of biological chemistry *277*, 45510-45517.

Ting, A.H., Jair, K.W., Schuebel, K.E., and Baylin, S.B. (2006). Differential requirement for DNA methyltransferase 1 in maintaining human cancer cell gene promoter hypermethylation. Cancer research *66*, 729-735.

Tran, R.K., Henikoff, J.G., Zilberman, D., Ditt, R.F., Jacobsen, S.E., and Henikoff, S. (2005). DNA methylation profiling identifies CG methylation clusters in Arabidopsis genes. Current biology : CB *15*, 154-159.

Tremblay, K.D., Duran, K.L., and Bartolomei, M.S. (1997). A 5' 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development. Molecular and cellular biology *17*, 4322-4329.

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., and Reuter, G. (1994). The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. The EMBO journal *13*, 3822-3831.

Tsukada, Y., and Zhang, Y. (2006). Purification of histone demethylases from HeLa cells. Methods *40*, 318-326.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H.R.*, et al.* (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes to cells : devoted to molecular & cellular mechanisms *11*, 805-814.

Tucker, K.L., Talbot, D., Lee, M.A., Leonhardt, H., and Jaenisch, R. (1996). Complementation of methylation deficiency in embryonic stem cells by a DNA

methyltransferase minigene. Proceedings of the National Academy of Sciences of the United States of America *93*, 12920-12925.

Ueda, J., Tachibana, M., Ikura, T., and Shinkai, Y. (2006). Zinc finger protein Wiz links G9a/GLP histone methyltransferases to the co-repressor molecule CtBP. The Journal of biological chemistry *281*, 20120-20128.

Umlauf, D., Goto, Y., Cao, R., Cerqueira, F., Wagschal, A., Zhang, Y., and Feil, R. (2004). Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. Nature genetics *36*, 1296-1300.

Vakoc, C.R., Mandat, S.A., Olenchock, B.A., and Blobel, G.A. (2005). Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. Molecular cell *19*, 381-391.

Vakoc, C.R., Sachdeva, M.M., Wang, H., and Blobel, G.A. (2006). Profile of histone lysine methylation across transcribed mammalian chromatin. Molecular and cellular biology *26*, 9185-9195.

Van den Wyngaert, I., Sprengel, J., Kass, S.U., and Luyten, W.H. (1998). Cloning and analysis of a novel human putative DNA methyltransferase. FEBS letters *426*, 283-289.

van Overveld, P.G., Lemmers, R.J., Sandkuijl, L.A., Enthoven, L., Winokur, S.T., Bakels, F., Padberg, G.W., van Ommen, G.J., Frants, R.R., and van der Maarel, S.M. (2003). Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. Nature genetics *35*, 315-317.

Vedadi, M., Barsyte-Lovejoy, D., Liu, F., Rival-Gervier, S., Allali-Hassani, A., Labrie, V., Wigle, T.J., Dimaggio, P.A., Wasney, G.A., Siarheyeva, A*., et al.* (2011). A chemical probe selectively inhibits G9a and GLP methyltransferase activity in cells. Nature chemical biology *7*, 566-574.

Venkatesh, S., Smolle, M., Li, H., Gogol, M.M., Saint, M., Kumar, S., Natarajan, K., and Workman, J.L. (2012). Set2 methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes. Nature *489*, 452-455.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A*., et al.* (2001). The sequence of the human genome. Science *291*, 1304-1351.

Vettese-Dadey, M., Grant, P.A., Hebbes, T.R., Crane- Robinson, C., Allis, C.D., and Workman, J.L. (1996). Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. The EMBO journal *15*, 2508-2518.

Wagschal, A., Sutherland, H.G., Woodfine, K., Henckel, A., Chebli, K., Schulz, R., Oakey, R.J., Bickmore, W.A., and Feil, R. (2008). G9a histone methyltransferase

contributes to imprinting in the mouse placenta. Molecular and cellular biology *28*, 1104-1113.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. Nature *431*, 873-878.

Wang, S.S., Smiraglia, D.J., Wu, Y.Z., Ghosh, S., Rader, J.S., Cho, K.R., Bonfiglio, T.A., Nayar, R., Plass, C., and Sherman, M.E. (2008). Identification of novel methylation markers in cervical cancer using restriction landmark genomic scanning. Cancer research *68*, 2489-2497.

Webster, K.E., O'Bryan, M.K., Fletcher, S., Crewther, P.E., Aapola, U., Craig, J., Harrison, D.K., Aung, H., Phutikanit, N., Lyle, R.*, et al.* (2005). Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. Proceedings of the National Academy of Sciences of the United States of America *102*, 4068-4073.

Weinstein, L.S., Xie, T., Qasem, A., Wang, J., and Chen, M. (2010). The role of GNAS and other imprinted genes in the development of obesity. Int J Obes (Lond) *34*, 6-17.

Weiss, T., Hergeth, S., Zeissler, U., Izzo, A., Tropberger, P., Zee, B.M., Dundr, M., Garcia, B.A., Daujat, S., and Schneider, R. (2010). Histone H1 variant-specific lysine methylation by G9a/KMT1C and Glp1/KMT1D. Epigenetics & chromatin *3*, 7.

Wu, M.Y., Tsai, T.F., and Beaudet, A.L. (2006). Deficiency of Rbbp1/Arid4a and Rbbp1l1/Arid4b alters epigenetic modifications and suppresses an imprinting defect in the PWS/AS domain. Genes & development *20*, 2859-2870.

Wutz, A., Smrzka, O.W., Schweifer, N., Schellander, K., Wagner, E.F., and Barlow, D.P. (1997). Imprinted expression of the Igf2r gene depends on an intronic CpG island. Nature *389*, 745-749.

Xin, Z., Tachibana, M., Guggiari, M., Heard, E., Shinkai, Y., and Wagstaff, J. (2003). Role of histone methyltransferase G9a in CpG methylation of the Prader-Willi syndrome imprinting center. The Journal of biological chemistry *278*, 14996-15000.

Xu, G.L., Bestor, T.H., Bourc'his, D., Hsieh, C.L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J.J., and Viegas-Pequignot, E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. Nature *402*, 187-191.

Xu, N., Donohoe, M.E., Silva, S.S., and Lee, J.T. (2007). Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. Nature genetics *39*, 1390-1396.

Yamashita, S., Takahashi, S., McDonell, N., Watanabe, N., Niwa, T., Hosoya, K., Tsujino, Y., Shirai, T., and Ushijima, T. (2008). Methylation silencing of transforming growth factor-beta receptor type II in rat prostate cancers. Cancer research *68*, 2112-2121.

Yang, X.J., and Seto, E. (2003). Collaborative spirit of histone deacetylases in regulating chromatin structure and gene expression. Current opinion in genetics & development *13*, 143-153.

Yatsuki, H., Joh, K., Higashimoto, K., Soejima, H., Arai, Y., Wang, Y., Hatada, I., Obata, Y., Morisaki, H., Zhang, Z.*, et al.* (2002). Domain regulation of imprinting cluster in Kip2/Lit1 subdomain on mouse chromosome 7F4/F5: large-scale DNA methylation analysis reveals that DMR-Lit1 is a putative imprinting control region. Genome research *12*, 1860-1870.

Yoder, J.A., and Bestor, T.H. (1998). A candidate mammalian DNA methyltransferase related to pmt1p of fission yeast. Human molecular genetics *7*, 279-284.

Yoder, J.A., Soman, N.S., Verdine, G.L., and Bestor, T.H. (1997). DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. Journal of molecular biology *270*, 385-395.

Yokochi, T., Poduch, K., Ryba, T., Lu, J., Hiratani, I., Tachibana, M., Shinkai, Y., and Gilbert, D.M. (2009). G9a selectively represses a class of late-replicating genes at the nuclear periphery. Proceedings of the National Academy of Sciences of the United States of America *106*, 19363-19368.

Yu, H., Lin, Q., Wang, Y., He, Y., Fu, S., Jiang, H., Yu, Y., Sun, S., Chen, Y., Shou, J.*, et al.* (2013). Inhibition of H3K9 methyltransferases G9a/GLP prevents ototoxicity and ongoing hair cell death. Cell death & disease *4*, e506.

Zentner, G.E., Tesar, P.J., and Scacheri, P.C. (2011). Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome research *21*, 1273-1283.

Zhang, X., Ercelen, S., Duportail, G., Schaub, E., Tikhonov, V., Slita, A., Zarubaev, V., Babak, V., and Mely, Y. (2008). Hydrophobically modified low molecular weight chitosans as efficient and nontoxic gene delivery vectors. The journal of gene medicine *10*, 527-539.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes & development *15*, 2343-2360.

Zhao, Y., and Epstein, R.J. (2008). Programmed genetic instability: a tumor-permissive mechanism for maintaining the evolvability of higher species through methylation-dependent mutation of DNA repair genes in the male germ line. Molecular biology and evolution *25*, 1737-1749.

Zhu, B., Zheng, Y., Pham, A.D., Mandal, S.S., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2005). Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. Molecular cell *20*, 601-611.

Zhu, J., and Yao, X. (2009). Use of DNA methylation for cancer detection: promises and challenges. The international journal of biochemistry & cell biology *41*, 147-154.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nature genetics *39*, 61-69.

Ziller, M.J., Gu, H., Muller, F., Donaghey, J., Tsai, L.T., Kohlbacher, O., De Jager, P.L., Rosen, E.D., Bennett, D.A., Bernstein, B.E.*, et al.* (2013). Charting a dynamic DNA methylation landscape of the human genome. Nature *500*, 477-481.

Zvetkova, I., Apedaile, A., Ramsahoye, B., Mermoud, J.E., Crompton, L.A., John, R., Feil, R., and Brockdorff, N. (2005). Global hypomethylation of the genome in XX embryonic stem cells. Nature genetics *37*, 1274-1279.